



EJBI 2010 ISSN 1801-5603

European Journal for Biomedical Informatics

Volume 6 (2010), Issue 2

An Official Journal of the European Federation of Medical Informatics

www.ejbi.eu



Aims and Scope

The European Journal for Biomedical Informatics reacts on the great European need to share the information in the multilingual and multicultural European area. The journal publishes peer-reviewed papers in English and other European languages simultaneously. This opens new possibilities for faster transfer of scientific-research pieces of knowledge to large international community of biomedical researchers, physicians, other health personnel and citizens.

The generally accepted translations of the English version of the paper to the official European languages and other European languages.

Cooperating journals

Methods of Information in Medicine
Lékař a technika

(language)

(English)
(Czech)

List of European languages

ISO 639-1 code

Albanian	sg
Armenian	hy
Azerbaijani	az
Belarusian	be
Bosnian	bs
Bulgarian	bg
Catalan	ca
Croatian	hr
Czech	cs
Danish	da
Dutch	nl
English	en
Estonian	et
Finnish	fi
French	fr
Georgian	ka
German	de
Greek	el
Hungarian	hu
Icelandic	is
Irish	ga
Italian	it
Kazakh	kk
Latvian	lv
Lithuanian	lt
Luxembourgish	lb
Macedonian	mk
Maltese	mt
Norwegian	no
Polish	pl
Portuguese	pt
Romanian, Moldavian, Moldovan	ro
Romansh	rm
Russian	ru
Serbian	sr
Slovak	sk
Slovenian	sl
Spanish	es
Swedish	sv
Turkish	tr
Ukrainian	uk



Editors and Management

Editor in Chief: Jana Zvárová, Czech Republic
Managing Editor: Petra Přečková, Czech Republic
Design PDF Version: Dana Vynikarová, Czech Republic
Sales and Marketing Manager: Libor Seidl, Czech Republic

Editorial Board

Ammenwerth, Elske (de)	Austria
Blobel, Bernd (de)	Germany
Bobrowski, Leon (pl)	Poland
Bureš, Vít (cs)	Czech Republic
Degoulet, Patrice (fr)	France
Dostálová, Taťjana (cs)	Czech Republic
Eryilmaz, Esat Nadir (tr)	Turkey
Hanzlíček, Petr (cs)	Czech Republic
Iversen, Irma (no)	Norway
Kern, Josipa (hr)	Croatia
Lukosevicius, Arunas (lt)	Lithuania
Mansmann, Ulrich (de)	Germany
Martin-Sanchez, Fernando (es)	Spain
Masic, Izet (bs)	Bosnia and Herzegovina
Mazura, Ivan (cs)	Czech Republic
McCullagh, Paul (en)	United Kingdom
Mihalas, George (ro)	Romania
Naszlady, Attila (hu)	Hungary
Nykänen, Pirkko (fi)	Finland
Paralič, Ján (sk)	Slovakia
Pisanelli, Domenico M. (it)	Italy
Sharp, Mary (ga)	Ireland
Sousa Pereira, Antonio (pt)	Portugal
Svačina, Štěpán	Czech Republic
Valenta, Zdeněk (cs)	Czech Republic
Vinarova, Jivka (bg)	Bulgaria
de Lusignan, Simon (en)	United Kingdom

Publisher

EuroMISE s.r.o.
 Paprskova 330/15
 CZ-14000 Praha 4
 Czech Republic
 EU VAT ID: Cz25666011

Office

EuroMISE s.r.o.
 Paprskova 330/15
 CZ-14000 Praha 4
 Czech Republic

Contact

Karel Zvára
 zvara@euromise.com,
 Tel: +420 226 228 904
 Fax: +420 241 712 990

Instructions to Authors

MANUSCRIPT SUBMISSIONS

European Journal of Biomedical Informatics (EJBI) is an international, peer-reviewed journal that publishes papers in the broad field of biomedical informatics. Manuscripts accepted for the electronic publication in EJBI are original contributions, reviews, brief reports, special communications, commentaries, and many other categories of papers. Due to special multilingual features of EJBI, these types of papers are published in English, but national language versions can be published simultaneously (see below).

Manuscripts should be sent electronically to the e-mail address: manuscripts@ejbi.org. Each manuscript should be submitted in MS Word, plain-text, HTML, TeX or LaTeX format and accompanied by image attachments and

- Identification Form: the name of the corresponding author with his contact address, phone number, fax number and e-mail address.
- Submission Requirement: the statement that the submitted paper has not been published in, nor has been submitted to, any other journal.
- Authorship Form: the covering letter signed by corresponding author, by which the author approves publication of the paper in the European Journal for Biomedical Informatics on behalf of all authors.
- Transfer of Copyright: All authors of the manuscript must have agreed to its publication and are responsible for its content and must also have agreed that the corresponding author has the authority to act on their behalf in all matters pertaining to publication of the manuscript. The corresponding author is responsible for informing the co-authors of the manuscript status throughout the submission, review, and publication process.

Submission Requirement, Authorship Form and Transfer of Copyright should be sent by fax +420 241 712 990 and by the surface mail to EuroMISE s.r.o., EJBI Editorial Office, Paprskova 330/15, 140 00 Prague 4, Czech Republic simultaneously.

Authors may supply a PDF file of the manuscript additionally. Submission of the PDF file only is not sufficient.

PUBLISHER

The European Journal of Biomedical Informatics is published by EuroMISE s.r.o., Prague, Czech Republic (tax identification number: CZ25666011). Papers published in the Internet version of EJBI may be republished in printed digest of EJBI.

LANGUAGE

All manuscripts should be written in English in an easily readable style. English version of the manuscript will be peer-reviewed. However, it is possible to submit another language version of the same paper additionally. The Internet version of the EJBI makes available all language versions submitted. It is also possible to publish the English version of the paper that has been already

published in the national journal, when the agreement of national publisher is obtained (copyright agreement). The English version follows a standard peer-review procedure.

ELECTRONIC FORM OF A MANUSCRIPT

Recommended formats of text files are: .DOC (Word version 6.0, 7.0, 97, 2000 or 2003), .RTF (Rich Text Format). The name of the text file should be identical with the first author's surname, e.g. NOVAK.DOC. Names of attached images, drawings or tables (i.e. figures and tables of the manuscript) should be composed of author's surname and figure or table number, e.g. NOVAKTAB3.BMP. Manuscripts along with attached files should be sent by e-mail to address manuscripts@ejbi.org.

ELECTRONIC IMAGE DOCUMENTATION

Figures and tables (bitmaps) should be submitted in BMP, GIF, TIF, PNG, EPS or JPEG format. JPEG files should have best compressive rate (10-20) for a good quality. All images (although included in the manuscript file) must be also submitted as separate files because the resolution of embedded images is often insufficient. Scanned images should have the resolution of at least 600x600dpi. Submissions that do not meet the Instructions for Authors will be returned.

ABBREVIATIONS AND NOMENCLATURE

Generally known abbreviations do not need to be explained. Abbreviations for symbols and expressions for terms should be spelled and they should be comprehensibly explained in brackets. Every such abbreviation should be explained only after its first occurrence in the text.

Measurements of length, height, weight and volume should be reported in metric units or their decimal multiples. All haematological and clinical chemistry measurements should be reported in the metric system in term of International System of Units (SI). Chemical substances should be described by their own systematic name or expression, medicines by general names. Commercial names of chemicals, medicines or technical innovations may be used after they have been defined by their scientific names. Radionuclides (radioisotopes) will be symbolised by the atomic number.

REVISED MANUSCRIPT SUBMISSION

When revision of a manuscript is requested, authors should return the revised version of their manuscript as soon as possible. The prompt action may ensure fast publication if a paper is finally accepted for publication in EJBI.

FINAL PROOFREADING

The Publisher will send the accepted paper to its author for final proofreading in the PDF format. The author may then correct printing errors only. No other changes or additions will be accepted. Author should send corrected and signed paper back to the Publisher by fax to +420 241 712 990 or by surface mail to: EuroMISE s.r.o., EJBI Editorial Office, Paprskova 330/15, 140 00 Prague, Czech Republic.

ORGANIZATION OF THE MANUSCRIPT

Title page. The first (title) page should contain the title of the paper, names and workplaces of all authors. Individual workplaces are necessary to be graphically differentiated (preferably by numeral as the upper index).

Abstracts and keywords. At the beginning the author puts an abstract and keywords. The abstract should be in the extent of 250-300 words. There should be 4 to 7 keywords, according to author's consideration, preferably from MeSH index.

Main text of the paper. General rules for writing manuscripts recommend use of simple and declarative sentences; avoid long sentences, in which meaning may be lost by complicated construction. All acronyms and abbreviations should be explained when they first appear in the text. The main text of the paper should follow the style of selected type of paper.

Acknowledgement. Acknowledgements, if any, should be given at the end of the paper, before bibliographic references.

References. References should be cited in the text by their index number according to the order of appearance in the manuscript. Each reference should be marked by its index number in square bracket corresponding to bibliography section. It is possible to include references to dissertation works and technical reports. It is obligatory to include information sufficient to look up referenced text.

Examples of references in bibliography section:

[1] Knaup P., Ammenwerth E., Brandner R., Brigl B., Fischer G., Garde S., Lang E., Pilgram R., Ruderich F., Singer R., Wolff A. C., Haux R., Kulikowski C.: Towards Clinical Bioinformatics: Advancing Genomic Medicine with Informatics Methods and Tools. *Methods Inf Med* 2004; 43, pp. 302-307

[2] Blobel B., Pharow P.: A Model-Driven Approach for the German Health Telematics Architectural Framework and the Related Security Infrastructure. In: *Connecting Medical Informatics and Bio-Informatics. Proceedings of MIE2005* (Eds. R. Engelbrecht, A. Geissbuhler, C. Lovis, G. Mihalas), Vol. 116, Amsterdam, IOS Press, 2005, pp. 391-396

[3] <http://www.infobiomed.org/>

Tables and Figures. Authors should use tables only to achieve concise presentation, or where the information cannot be given satisfactory in another way. Tables should be numbered consecutively using Arabic numerals and should be referred to in the text by numbers. Each table should have an explanatory caption that should be as concise as possible. Figures should be clear, easy to read and of a good quality. Styles and fonts should match those in the main body of the paper. All figures must be mentioned in the text in consecutive order and should be numbered with Arabic numerals.

Authors should indicate precisely in the main text where tables and figures should be inserted, if these elements are given only separately or at the end in the original version of the manuscript.

Content

English version

- en 91 Editorial
Jana Zvárová
- en 92 - 94 Task Force of the European Federation of Medical Informatics Journals: Background, Rationale and Purpose
Izet Masic, Josipa Kern, Jana Zvárová, Simon de Lusignan
- en 95 - 102 Robust Image Analysis of Faces for Genetic Applications
Jan Kalina
- en 103 - 109 Reliability of Composite Dichotomous Measurements
Patrícia Martinková, Karel Zvára

Czech version

- cs 91 Předmluva
Jana Zvárová
- cs 92 - 94 Pracovní skupina zaměřená na časopisy Evropské federace pro medicínskou informatiku: pozadí, odůvodnění a účel
Izet Masic, Josipa Kern, Jana Zvárová, Simon de Lusignan
- cs 95 - 102 Robustní analýza obrazu obličeje pro genetické aplikace
Jan Kalina
- cs 103 - 109 Spolehlivost složených dichotomních měření
Patrícia Martinková, Karel Zvára

Croatian version

- hr 3 - 5 Akcijski tim za podizanje učinkovitosti EFMI-evih časopisa: Osnova, logika i svrha
Izet Masic, Josipa Kern, Jana Zvárová, Simon de Lusignan



EJBI 2010 ISSN 1801-5603

European Journal for Biomedical Informatics

Volume 6 (2010), Issue 2
English version

An Official Journal of the European Federation of Medical Informatics

www.ejbi.eu



Content

English version

- en 91 Editorial
Jana Zvárová
- en 92 - 94 Task Force of the European Federation of Medical Informatics Journals: Background, Rationale and Purpose
Izet Masic, Josipa Kern, Jana Zvárová, Simon de Lusignan
- en 95 - 102 Robust Image Analysis of Faces for Genetic Applications
Jan Kalina
- en 103 - 109 Reliability of Composite Dichotomous Measurements
Patricia Martinková, Karel Zvára

Editorial - EFMI and active medical informatics journals in Europe

Biomedical informatics is a burgeoning field, with important applications and implications throughout the biomedical world and healthcare delivery. The European Journal of Biomedical Informatics (EJBI) is reacting on the great European need to share the information in the multilingual and multicultural European area.

In the year 2010 the EJBI became an official journal of the European Federation

for Medical Informatics (EFMI). With these changes the journal will stress the role of information technologies and national journals supporting wide communication in many European languages.

The role of national and international medical informatics European journals is very important. In the first paper of the second issue of EJBI 2010 the role of EFMI and European medical informatics journals is described in details. The other two

papers in this issue focus on robust image analysis of faces for genetic applications and on reliability of continuous dichotomous measurements.

Jana Zvárová
Editor-in-Chief

Task Force of the European Federation of Medical Informatics Journals: Background, Rationale and Purpose

Izet Masic¹, Josipa Kern², Jana Zvárová³, Simon de Lusignan⁴

¹Society for Medical Informatics of Bosnia and Herzegovina, Sarajevo, Bosnia and Herzegovina,

²Croatian Society for Medical Informatics, Zagreb, Croatia,

³Czech Society for Biomedical Engineering and Medical Informatics, Prague, Czech Republic,

⁴Division of Community Health Sciences, St George's University of London, UK

Reprinted from: *Medical Informatics in a United and Healthy Europe*, K.P. Adlassnig et al. (Eds), IOS Press, 2009, 946-950, © 2009 European Federation for Medical Informatics.

Summary

The paper presents an analysis of how EFMI disseminates new knowledge and of the active medical informatics journals in EFMI member countries, which was carried out as an outcome of the EFMI Council meeting in London in 2008. The analysis identifies eight active major informatics journals and several other publications. Most are subscription-based and published at least quarterly. There is a possibility for the editors to meet regularly and form a community of practice with the aim of further improving their effectiveness in disseminating new knowledge and best practice in medical informatics. It is feasible to share expertise and it may be possible to harmonise several aspects of preparation and submission of manuscripts so that some of the identified barriers in publishing are reduced.

1. Introduction

The European Federation for Medical Informatics (EFMI) is a non-profit scientific and professional organization established in 1976 that today consists of 31 national members [1], [2], [3], [4]. All European countries are entitled to be represented in EFMI by an appropriate national medical informatics society. The term '*medical informatics*' is used to include the whole spectrum of health informatics and all its sub-disciplines and allied disciplines. The Medical Subject Headings (MeSH) defines medical informatics as [4], [5]: '*The field of information science concerned with the analysis and dissemination of medical data through the application of computers to various aspects of health care and medicine.*' This definition was introduced in 1987 as a MeSH term.

EFMI operates with a minimum of bureaucratic overhead with each national society supporting the federation by sending a representative to participate in the council. English has been adopted as the official language, although simultaneous translation is often provided for congresses in non-English speaking countries [6]. EFMI has 16 working groups: Smart Cards in Health Care 'CARDS'; Social Security and Welfare; Education in Health Informatics 'EDU'; Electronic Health Records 'EHR'; Evaluation and Assessment of Health Information Systems 'EVAL'; Informatics for Disabled People and Rehabilitation 'IDR'; Information Planning and Modelling in Health Care 'IPAM'; Libre/Free and Open Source Software 'LIFOSS'; Case-mix and Severity of Cases; Medical Informatics in Transition Countries 'MICIT'; Medical Imaging Processing 'MIP'; Natural Language Understanding 'NLU'; Nursing Informatics in Europe 'NURSIE'; Human and Organisational Factors of Medical Informatics 'HOFMI'; Primary Care Informatics 'PCI'; Safety, Security and Ethics 'SSE'; and Traceability 'TRA'.

To date, EFMI has organized 21 European congresses of medical informatics, including the Medical Informatics Europe (MIE) congresses in Cambridge (1978), Berlin (1979), Oslo (1988), Glasgow (1990), Vienna (1991), Jerusalem (1993), Lisbon (1994), Copenhagen (1996), Thessaloniki (1997), Ljubljana (1999), Hannover (2000), Budapest (2002), St. Malo (2003), Geneva (2005), Maastricht (2006), Gothenburg (2008) and Sarajevo (2009). The MIE 2011 conference will be held in Oslo. Traditionally, MIE conferences were not held in the year in

which the International Medical Informatics Association (IMIA) organises its triennial conference MEDINFO.1 EFMI has also started a new series of meetings: the Special Topic Conferences (STCs). The STC concept has the following components: a) organization should be done by a member society possibly in combination with its annual meeting; b) EFMI Council meeting should be the integral part of STC; c) topic of the conference should be defined by the organizing member society; d) considering the topic, the relevant EFMI working groups should be engaged; e) contributions should be mostly by invitation; f) STCs are small 2-day conferences with 100+ participants.

The first conference STC took place in Bucharest/Romania 2001, then in Nicosia/Cyprus 2002, Rome/Italy 2003, Munich/Germany 2004, Athens/Greece 2005, Timisoara/Romania 2006, Brijuni/Croatia 2007, London/United Kingdom 2008, and Antalya/Turkey 2009. In 2010, the STC will take place in Reykjavik/Iceland, and in 2011 in Ptuj/Slovenia. The proceedings of these conferences were usually published by Springer in the series '*Lecture Notes in Medical Informatics*' and by IOS Press in the series '*Studies in Health Technologies and Informatics*'. A selection of the best papers from the MIE conferences were published in a special volume of the International Journal of Medical Informatics and will be published in the Methods of Information in Medicine, while the special proceedings of STCs is published by IOS Press as collection of peer reviewed articles.

2. Review of Medical Informatics Journals in Europe

Medical informatics scientific and research production in Europe is growing both in quantity and quality. Promoting high-quality research is a major goal of EFMI [1], [4], [6]. EFMI has two highly respected official general journals, the International Journal of Medical Informatics (former title: International Journal of Bio-Medical Computing), currently edited by Charles Safran and Jan Talmon, and Methods of Information in Medicine, currently edited by Reinhold Haux. EFMI also publishes several sub-speciality official journals covering the spectrum of medical informatics sub-disciplines. The time-honoured method to disseminate high-quality scientific research is via the proceedings of the MIE annual conference. EFMI also plays a major role in education and in the harmonization of medical informatics sub-disciplines through the organization of STCs. Both MIE and STC proceedings are peer reviewed and published by IOS Press (Amsterdam, The Netherlands). Usually, MIE proceedings are indexed in Medline/PubMed. Most medical informatics journals are published in local languages, but many of them also incorporate English summaries or contain combined articles published in domestic and English language. Altogether, EFMI provides highly effective means for worldwide dissemination of medical informatics research conducted in Europe. However, in addition to the proceedings, journal editors have the opportunity to do more to disseminate knowledge and increase the understanding of our discipline.

3. Analysis of Publishing Activities within EFMI Members

Most of the information about the EFMI can be found on the official website (www.efmi.org).

Furthermore, most of national societies have their own websites providing useful information about publications and meetings in EFMI member countries. Our analysis is based on the facts generated from EFMI database stored on EFMI website and collected from national

representatives of EFMI member countries.

A search of Medline with the search terms 'EFMI' and 'European Federation for Medical Informatics' returns only 21 papers – 18 in medical informatics related journals: Informatics in Primary Care (6), International Journal of Medical Informatics (3), Methods of Information in Medicine (2), Studies in Health Technologies (6); and 3 in noninformatics/non-EFMI journals. Authors of the papers are: Blobel (2), De Lusignan (6), France (1), Horsch (1), McKeon (1), Masic (2), Michalas (2), Naszlady (1), Tallberg (1), Talmon (1), Trpisovský (1), van Goor (1), and Zvarova (1). Medline search using the term 'IMIA' was performed (for comparison) and 150 papers were found; the search term 'AMIA' resulted in 244 papers found. Most of EFMI papers are related to the activities of working groups – PCI is the most prominent [5], [6] – and relatively few initiatives at the level of the federation. Detailed analysis of EFMI website showed that only 14 of 31 members have their official scientific or professional medical informatics journal or newsletter listed (Table 1). Medical informatics journals have their articles published in English with three published in their national language. There are three journals published only as e-journals.

The frequency of ISSN registered publications varies: one journal is published biannually, two quarterly, one every two months, five journals are published 2-3 times per year, two journals monthly and one journal nine times per year. Medical informatics is the principal subject of all of the journals, but several of them focus on other sub-disciplines and allied disciplines. Most of the medical informatics journals are international publications, though three are restricted to national circulation. Two journals are financed by ministry of health and/or other public institutions, while others use subscription as the way of financing.

4. Mission statements of task force of EFMI journal editors

Editor-in-Chiefs of medical informatics journals are invited to create a joint mission statement and set out the objective and

purpose of all EFMI journals. We believe that the joint statement will foster the dissemination of scientific knowledge and increase our understanding of medical informatics as a discipline. Additionally, we intend to produce and issue a core document stating the fundamental principles upon which all medical informatics journals editors should agree. The common goals will be identified and agreed on the scope and standard of papers that should be published in EFMI medical informatics journals. The reasons for setting out the statement are:

- To promote editorial excellence by enforcing the use of standards and guidelines (International Committee of Medical Journals. Editors guidelines; Committee of Publication Ethics; other relevant documents [7]);
- To improve the quality of scientific studies published in EFMI medical informatics journals by imposing same standards for evaluating studies in medical informatics (guidelines already accepted as official EFMI document, but there were no further steps taken to ensure their use);
- To promote scientific publishing excellence by enforcing standards of paper formatting ('camera-ready' manuscript format; Vancouver referencing style);
- To improve diffusion of scientific knowledge through the medical/health informatics area (recognition and diffusion of EFMI research, education, clinical practice guidelines in EU countries should be promoted);
- To increase collaboration among EFMI medical informatics journal editors and schedule regular meetings of the Task Force;
- To explore the potential for using shared pool of reviewers, sharing information on review process and sharing expertise: technical editorial information, experiences, initiatives, publishing resources and tools (such as the open source Open Journal System or commercial/publisher solutions);
- To encourage articles not suitable for medical informatics journals to be published elsewhere and support the communication with editors of such journals;

Tab. 1. - EFMI member countries medical informatics journals.

EFMI country	Journal title	Publisher	Editor-in-Chief	ISSN	Esta. year	Issues per year	Submi-ssion	Web site	Language	Field	Peer reviewed	Financed by
Austria	Artificial Intelligence in Medicine	Elsevier	Klaus-Peter Adlassnig	1386-5056	1989	9	email	yes	English	MI, other	internat.	subscription
Bosnia and Herzegovina	Acta Informatica Medica	AVICENA, Sarajevo	Izet Masic	0353-8109	1993	4	email	yes	English	MI, other	internat.	subscription
Croatia	Bilten HDMI (Bulletin of the CSMI)	CSMI	Kristina Fišter	1330-0253	1992	2	email	yes	Croatian	MI, other	no	membership in the society
	Medicinska informatika	SBMILI	Josipa Kern	1330-1799	1993	2			English, Croatian	MI	national	ministry of science, publicity
Czech Republic	Lékař a Technika	EuroMISE	Karel Roubik	0301-5491	1970	4	email	yes	Czech	MI, biomedical engineering	national	public institutions, subscription
	European Journal for Biomedical Informatics	DSMI	Jana Zvarova	1801-5603	2005	2	email	yes	English, other	Biomedical informatics	internat.	other
Denmark	Newsletter	Springer-Verlag	-	-	-	-	-	yes	Danish, English	MI	national	public institutions
France	Informatique et Sante	Urban&	Patrice Degoulet	-	-	2	email	yes	French, English	MI	national	public institutions, subscription
Germany	Biometrie, Informatik und Epidemiologie in Medizin und Biologie	Fischer Verlag	-	-	-	-	email	yes	German	MI, other	national	public institutions, subscription
Ireland	Newsletter	-	-	-	-	-	-	-	-	-	-	-
Israel	ILAMI journal and The Health Screen	SMIA	-	-	-	2	-	-	-	MI, other	-	-
Slovenia	Informatica Medica Slovenica	SEIS	Gaj Vidmar	1318-2129	1993	2	email	yes	English, Slovenian, other	MI, biostatistics, other	internat., national	MI society (SMIA)
Spain	I+S	Elsevier	-	-	-	6	-	-	-	MI, other	-	-
Sweden	Computer Methods and Programs in Biomedicine	UAMI	Torny Groth	0169-2607	1985	12	paper, email, online	yes	English	MI	internat.	subscriptions
Ukraine	Clinical Informatics and Telemedicine	Radcliffe	Oleg Yu Mayorov	1812-7231	2003	2-3	email	yes	Ukrainian, English	MI, other	internat.	public institution
United Kingdom	Informatics in Primary Care	Publishing	Simon de Lusignan	1476-0320	1993	4	paper, online	yes	English	PHC, Clinical Informatics	internat.	PHCG BCS, subscriptions
EFMI	International Journal of Medical Informatics	Elsevier	Charles Safran, Jan Talmon	1386-5056	1961	12	paper, online	yes	English	MI	internat.	subscriptions
	Methods of Information in Medicine	Schattauer	Reinhold Haux	0026-1270	1962	6	paper, online	yes	English	MI	internat.	subscriptions

- To provide a common voice when issues of common interest arise and enhance collaboration between national societies and EFMI bodies;
- To promote European initiatives in stimulating publications and top-quality research.

Acknowledgment

The authors would like to acknowledge the support of Jan Talmon from Maastricht University, Care and Public Health Research Institute, The Netherlands.

References

- [1] Masic I, Novo A.: History of medical informatics in Bosnia and Herzegovina. Acta Informatica Medica. 2007;15(1):49-61.
- [2] Collen M. F.: Fifty years in medical informatics. IMIA Yearbook of Medical Informatics. Methods of Information in Medicine. 2006;45(Suppl. 1):S174-9.
- [3] Zvárová J.: Symposium on "Biomedical Informatics and Biomedical Statistics Education". Methods of Information in Medicine. 2006;45(3):281-2.
- [4] Zvárová J.: Biomedical Informatics Research and Education in the EuroMISE Center. IMIA Yearbook of Medical Informatics. Methods of Information in Medicine. 2006;45(Suppl. 1):S166-73.
- [5] de Lusignan S.: Introduction to the Primary Care Informatics Working Group of the European Federation for Medical Informatics (EFMI). Informatics in Primary Care. 2003;11(3):175-6.
- [6] de Lusignan S., Teasdale S.: Achieving benefit for patients in primary care informatics: The report of an international consensus workshop at MedInfo 2007. Informatics in Primary Care. 2007;15(4):255-61.
- [7] Miller R. A., Groth T, Hasman A, Safran C., Shortliffe E. H., Haux R., McGray A. T.: On exemplary scientific conduct regarding submission of manuscripts to biomedical informatics journals. Methods of Information in Medicine. 2006;45(1):1-3.

Contact

Izet Masic
 Medical Faculty
 University of Sarajevo
 Cekalusa 90/4
 71000 Sarajevo
 Bosnia and Herzegovina
 e-mail: imasic@lol.ba

Robust Image Analysis of Faces for Genetic Applications

Jan Kalina¹

¹Centre of Biomedical Informatics, Institute of Computer Science AS CR, Prague, Czech Republic

Summary

This paper is devoted to automatic localization of objects (eyes, mouth) in two-dimensional (2D) grey scale images of faces. Motivated by a practical problem in human genetics, the output of the localization of objects in the given database of images is needed for further tasks in the genetic research. A robust filter is applied on the image to ensure denoising. Templates are used as the main method. The mouth and both eyes are localized jointly using the weighted Pearson product-moment correlation coefficient or its robust analogy based on robust regression methods. In the database with 212 images of faces the method allows to locate the mouth and eyes correctly in 100 % of cases. Also the robust correlation coefficient based on the least weighted squares regression localizes the mouth and both eyes in 100 % of images of the given database. Robustness aspects of the method are examined with respect to rotation, noise, occlusion and asymmetry in the image. The joint localization of the mouth and both eyes produces the method invariant to rotation of any degree. This work is tailor made for the given images with expected usage of the methods in genetic applications.

Keywords: object localization, template matching, eye or mouth detection, robust correlation analysis, image denoising

1. Introduction

The primary motivation for this work is the automatic location of landmarks in images of human faces motivated by genetic research. We are working with a database of images which come from the Institute of Human Genetics, University of Duisburg-Essen, Germany (projects BO 1955/2-1 and WU 314/2-1 of the German Research Council). This database contains 212 grey value images of the size 192 times 256 pixels, each image corresponding to a different person. The faces have about the

same size but are rotated in the plane by small angles. Our aim is to find a robust solution with respect to rotation, occlusion, noise in the image and asymmetry of the face, while allowing for a clear interpretation of the method.

Template matching is a tailor made method for object detection in grey scale images. A template is a model, a typical form, an ideal object. It is placed on every possible position in the image and the similarity is measured between the template and each part of the image, namely the grey value of each pixel of the template is compared with the grey value of the corresponding pixel of the image. [1] gives references on template matching applied to face detection and face recognition. Nevertheless standard image analysis procedures are formed as a cascade of extremely simple classifiers [2]. On the other hand ignoring the usual procedures of dimension reduction and feature extraction can ensure a clear interpretation and allows theoretical robustness considerations [3].

[4] is an example of template matching applied to the data set of irises of 64 persons with the task to assign a new iris to the correct person from the training database. The method maximizes the mutual information as the similarity measure between the template and the red channel of the color image. [5] considers templates as elements of a tree-structured hierarchy. More sophisticated algorithms of image analysis combine ad hoc methods of mathematics, statistics and informatics [2] accenting high computational speed rather than convenient theoretical properties [3]. Recent papers on image analysis replace pixels by features or patches. Patches [6] are homogeneous areas of pixels and the patch-based approach analyzes the image as a set of individual patches. On the other hand features typically correspond to edges or objects with heterogeneity or

discontinuity. [7] was the first paper reducing the dimension by replacing pixel intensities by a feature set. [8] and [9] are recent works on human or face detection by extracting features with the aim to propose methods possibly robust to illumination changes, different pose or facial expressions. [10] studies variance matrices of features, because variance is robust to illumination changes.

We work with the database containing images which are matrices with the size 192*256 pixels. A grey value in the interval [0,1] corresponds to each pixel, where low values are black and large values white. Images are photographed under the same conditions, with the person sitting straight against the camera looking straight at it. The Institute of Human Genetics tried to have the images standardized as much as possible. For example there are no images with closed eyes, hair over the face covering the eyes or other nuisance effects. Still the faces in the images happen to be rotated by a small angle. The eyes are not in a perfectly horizontal position in such images. The database does not include images with a three-dimensional rotation (a different pose).

The Institute of Human Genetics is working on interesting problems in the genetic research using images of faces [11]. The ambitions of the research are to classify automatically genetic syndromes from a picture of a face; to examine the connection between the genetic code and the size and shape of facial features; and also to visualize a face based only on its biometric measures [12]. Images of patients can be classified to one of 10 groups according to a genetic malformation deforming the face. For different syndromes the success rate lies between 75 % and 80 %, which is considered remarkably successful.

Locating the landmarks is always the first step of all such procedures, however not the primary goal of the study. The landmarks are prominent parts of the face, for example the corners of the eyes and the mouth, the midpoint of the top and the bottom edges of the lips or significant points of the nostrils and eyebrows. The team of genetics researchers uses two approaches to locate 40 landmarks in each face. One possibility is the manual identification, which is carefully and accurately performed by an anthropologist trained in this field. As the second approach the Institute uses an automatic method, based on the algorithm [13], which will be now described.

The algorithm starts by manual location of the set of 40 landmarks in a training set of 83 images of faces. These landmarks are called fiducial points and they together are placed on all positions in the image as one large template retaining fixed distances between the landmarks. Two-dimensional Gabor wavelet transformations with different values of the two-dimensional scale parameter are applied on all the training images and also on a new image in which the landmarks are to be located. The jets (Gabor wavelet coefficients) in each landmark of the training image and the jets in the corresponding pixels of the new image are compared. We can understand the jet of each of the training images as a (multi-dimensional) template. The correlation coefficient between the vectors of wavelet coefficients (or only their magnitudes) is computed and their sum over all 40 landmarks is used as the similarity measure between the training image and the new image. However, such an approach turns out to be vulnerable to small rotations of the face.

The aim of our work is to search for the mouth and eyes in images of faces using templates. We propose an algorithm for localizing each of the eyes separately and also jointly both eyes and the mouth using templates. We inspect robustness properties of the described method, for example the robustness to noise or rotation of the image. The paper has the following structure. Chapter 2 describes an initial denoising of images obtained by applying a robust filter. Templates are used as the main method to localize the eyes

(Chapter 3) and jointly both eyes and the mouth (Chapter 4). Robustness aspects of the method are examined and robust analogies of the correlation coefficient are inspected in Chapter 5.

2. Methods

Our approach begins with image denoising. We describe an approach for locating the eyes in the image and then the joint localization of both eyes and the mouth. Finally we compare different methods of robust correlation analysis for the same task of localizing both eyes and the mouth jointly. We must admit that we lose the advantage of feature-based methods described in Chapter 1, which allow to work with different sizes of the images. Nevertheless our approach has a clear interpretation and we are able to obtain a rotation-invariant approach. The results are presented in Chapter 3.

2.1 Image denoising

Denoising, filtering or robustification is a transformation often used to remove noise from images. We summarize arguments in favor of image denoising from references and describe our method based on the least trimmed squares or least weighted squares estimators. The motivation for such procedure is to remove noise from images while retaining the facial features well recognizable. Another artefact in our database of images is the reflection of light bulbs at different positions in the eyes of persons. This nuisance effect caused by the method of photographing at the Institute of Human Genetics is also removed by the image denoising. We describe these very robust estimators for a linear regression context, because they are used also later in Chapter 2.4 to define a robust correlation coefficient.

[14] describes filters (two-dimensional operators) for denoising and prefers the trimmed mean and other L-estimators [15] to the median. [16] proposes an M-estimator correlation coefficient, gives theoretical arguments in favor of the combining the robustness and efficiency for Gaussian white noise and applies it to the image analysis of templates. Denoising is applicable also to molecular genetics images [17] or alternatively [18] based on robust statistics.

We proceed to the definition of the robust statistical estimators, which are used for the image denoising in our work. Let us consider the linear regression model in the form

$$Y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} + e_i, i = 1, 2, \dots, n, \quad (1)$$

or in the matrix notation $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}$. The least weighted squares (LWS) regression estimator proposed by [19] is one of robust regression methods with a high breakdown point. It requires the specification of the sizes of the non-negative weights w_1, w_2, \dots, w_n . However these are assigned to particular data observations only after a permutation, which is determined in an implicit way during the computation of the estimator. This permutation depends on residuals $u_1(\mathbf{b}), \dots, u_n(\mathbf{b})$ corresponding to a particular value $\mathbf{b} = (b_0, b_1, \dots, b_p)^T$ of the estimator of the vector parameter $\boldsymbol{\beta}$, where

$$u_i(\mathbf{b}) = Y_i - b_0 - b_1 x_{i1} - \dots - b_p x_{ip}, i = 1, 2, \dots, n. \quad (2)$$

Typically the weights w_1, w_2, \dots, w_n are chosen as a non-increasing sequence. Denoting arranged squares residuals as

$$u_{(1)}^2(\mathbf{b}) \leq u_{(2)}^2(\mathbf{b}) \leq \dots \leq u_{(n)}^2(\mathbf{b}), \quad (3)$$

the estimator of $\boldsymbol{\beta}$ is defined as

$$\hat{\mathbf{b}}_{LWS} = \underset{\mathbf{b}}{\operatorname{argmin}} \sum_{i=1}^n w_i u_{(i)}^2(\mathbf{b}).$$

A popular choice [20] is to use linearly decreasing weights $w_i = 1 - (i-1)/n, i = 1, \dots, n$; another possibility is a two-stage procedure of [21] for the computation of adaptive weights, allowing to determine also the sizes of the weights automatically.

The least weighted squares regression combines the robustness and efficiency [21]. A fast approximative algorithm for computing the LWS can be obtained as a weighted analogy of [22]. [23] proposed the least trimmed squares (LTS) regression, which is a special case of the LWS regression with weights equal to 1 or 0 only. It must be specified that exactly h of the data points will have the weight equal to 1.

We apply these robust estimators in the context of image denoising. Their advantage is the high breakdown point [15] ensuring high resistance against outliers. Instead of the linear regression, only the location model is relevant. The least median of squares (LMS) estimator [23] is equivalent to the mean of the shortest half of the data. The LTS estimator corresponds to the mean of such half of the data (or a group of some $h < n$ observations), which has the smallest variance. The LWS estimator is the weighted mean of the data with such permutation of the weights yielding the smallest weighted variance of the data [20].

For each pixel we take the grey values from its circular neighborhood and compute the least median of squares (LMS), least trimmed squares (LTS) or least weighted squares (LWS) estimator. The information about the coordinates is lost. Poor results are obtained with the median because it removes contrast and the resulting image is rather greyish. The performance of the LTS and LWS estimators is reliable with a small radius of the circular neighborhood, however not strongly influenced by the choice of h for the LTS or by the choice of weights for the LWS estimator.

Therefore we use the LTS filter on each image in the database. For each pixel we consider its four direct neighbors and compute the LTS estimator with $h = 3$, which is computed as the mean of the shortest triplet obtained from the values arranged in ascending order. This removal of possible extreme outliers removes also the light reflected in eyes to some extent. To examine the effect of denoising, we examined the residuals of this transformations. Large absolute values of the residuals indicate a larger local effect of denoising. This effect of the transformation is remarkably attained at edges in the image, such as between the hair or shirt and the background or at the boundary of the nostrils.

2.2 Locating eyes

We use template matching to search for the right and left eye in each image separately. It is popular to use the mean of eyes as the template for the eye (compare [24], [25]). Here we construct a set of 6 eye

templates for the right eye and their reflections in the axial symmetry are used for the other eye. The templates are obtained as the mean of real eyes of the same size of different persons; one of them is shown in Figure 1. The templates have different sizes between 26×28 pixels and 36×30 pixels. The Pearson product-moment correlation coefficient (further called simply correlation coefficient) or the weighted Pearson product-moment correlation coefficient r_w (further called weighted correlation coefficient) with suitable weights is used as the similarity measure between the template and the image.



Fig. 1. An eye template.

All of the templates contain eyebrows, which could possibly complicate the recognition. Nevertheless the area of eyebrows will be down-weighted by the

weighted correlation coefficient with radial weights. Figure 1 shows such an eye template, which yields the best performance in locating the right eye; in this case this template alone was used together with the weighted correlation coefficient with radial weights.

We compute the correlation coefficient between one of the images and the eye template from Figure 1. The correlation coefficient is shown in Figure 2, where black areas have a large value of the correlation coefficient. These areas include both eyes and also the mouth or parts of the hair. It happens for example that the right eye of a particular person has a larger correlation with one of the left eye templates than with any of the right eye templates. We use this phenomenon in the algorithm in the following way. All the twelve templates are placed on every possible position in the image. We do not distinguish between left and right eye templates. Firstly the area with the largest correlation coefficient over all templates is found. That is one candidate eye. Now the whole image without this candidate eye and its nearest neighborhood is considered. Again the area with the largest correlation with any of the eye templates is found. That is the other candidate eye.



Fig. 2. The correlation coefficient of a given image and the template for the right eye was computed. Areas of the image with a large correlation coefficient are shown black.

2.3 Locating eyes and mouth

To localize jointly the mouth and both eyes, we start with a region which is candidate to contain the mouth and search for the eyes in a certain region above the candidate mouth. We use 6 templates for the right eye, their reflections in the axial symmetry and 7 symmetric mouth templates. Each mouth template is used to find several areas, which have the weighted correlation coefficient with any of the mouth templates above a certain threshold. This can be different in different images. In any case we search for at least three such areas, which are not direct neighbors of each other. These areas are considered to be candidate mouths. First we select areas with the weighted correlation coefficient r_w above the value of 80 % of the maximal r_w , which was attained in the particular image. This boundary can be decreased if necessary.

Different mouth templates yield different candidate mouths. Then the eyes are searched for in such areas, which correspond to particular candidate mouths. These areas range from 20 to 55 pixels above candidate mouths and are considered to be candidate areas for eyes. We place each of the 6 eye templates and their reflections in the axial symmetry on every possible position in these candidate eyes and compute the weighted correlation coefficient between the template and the corresponding region of the same size. At the same time the area with the expected position of the eyes is divided to two parts along the virtual vertical line (Figure 2), which bisects the mouth. All eye templates and also their versions reflected in the axial symmetry are used in both parts of the candidate area for eyes. We search for the area with the largest weighted correlation coefficient with any of the templates in the left and right parts separately. This approach is repeated for several candidate mouths.

Moreover we use a condition on the mutual distance of both eyes in the limit between 25 and 42 pixels. These bounds are obtained empirically and real eyes fulfill the condition also in images with a slightly modified size or rotated by a small angle.

For a given candidate mouth we find a candidate area for the right and left eye. If

these do not fulfill the condition on the distance between eyes, we continue searching for such area, which has the largest weighted correlation coefficient with any eye template among all remaining areas in the candidate area for eyes. Such area with the midpoint in the candidate area is one candidate eye, while the other candidate eye has the largest weighted correlation coefficient with any of the templates and must have the midpoint in the opposite part of the candidate area. Again the condition on the eye distance is checked and if needed these steps are repeated, until two candidate eyes are found, which satisfy the condition. This method distinctively improves the results for rotated images (Chapter 3).

The basic idea is to add three weighted correlation coefficients corresponding to the mouth and both eyes. Let us start with one of pixels, which is the midpoint with a large weighted correlation coefficient with any mouth templates. Let us denote the largest value among these weighted correlation coefficients between the area and any of the templates as r_{w1} . We place all eye templates on every such possible position with the midpoint in the left part of the candidate area for eyes. The eyes are searched for only in the left part of this area and let us denote the largest weighted correlation coefficient over all possible eye templates by r_{w2} . In a similar manner we place all eye templates to the right part of the candidate area and denote the largest weighted correlation coefficient by r_{w3} .

Let us consider the coefficient

$$r_w^* = (r_{w1} + r_{w2} + r_{w3}) \cdot I(r_{w1} > 0) \cdot I(r_{w2} > 0) \cdot I(r_{w3} > 0), \tag{4}$$

where 1 denotes an indicator. (4) ignores negative values of the weighted correlation coefficient. The largest value of r_w^* is computed for different candidate mouths.

2.4 Robust correlation

We apply a robust version of the correlation coefficient to the joint localization of both eyes and the mouth in the images. The method is based on the robust estimators for the linear regression context (Chapter 2). Trimming away some pixels corresponds to the idea that some pixels are irrelevant and ignoring large portions of pixels may resemble patch-based approaches [26].

We compute the LWS-based correlation coefficient between the image and the template by transforming the matrices to vectors, computing the LWS regression of the image against the template and finally computing the weighted correlation coefficient with the weights determined by the LWS. For the LWS we consider linearly decreasing weights and the adaptive weights [21]. The LTS-based correlation coefficient is defined in an analogous way and we choose to trim away 20 % of the pixels ($h = 0.8n$).

Tab. 1. Percentages of correct results for the localization of eyes using 6 templates and the joint localization of mouth and eyes using 7 mouth templates and 6 eye templates. The weighted Pearson product-moment correlation coefficient is used with radial weights. The original images were also modified by making the size smaller by 10 % or by rotating the image by $+10^\circ$ or -10° .

Templates	Original image	Smaller image	Rotated image
Eyes, r	1.00	0.80	0.50
Eyes, r_w	1.00	0.97	0.86
Eyes and mouth, r	1.00	0.99	1.00
Eyes and mouth, r_w	1.00	1.00	1.00
Eyes and mouth (relaxed), r	1.00	0.99	0.92
Eyes and mouth (relaxed), r_w	1.00	0.99	0.87

3. Results

3.1 Locating eyes

The set of 6 eye templates (Chapter 2.2) allows both eyes to be localized correctly in 100 % of images of the database, using the product-moment correlation coefficient as the similarity measure between the image and the template. This means that the output of the localization corresponds to true eyes in all 100 % of images of the database. The 100 % correct results are obtained also with the weighted correlation coefficient with radial weights defined to be inversely proportional to the distance of each pixel from the midpoint of the template and yield better results in terms of robustness properties (Chapter 4).

Table 1 presents the results of locating both eyes separately by 6 eye templates and their mirror reflections. Further it gives results described later. Results obtained with r and r_w with radial weights are compared.

3.2 Locating eyes and the mouth

The method (4) for the joint localization of both eyes and the mouth (Chapter 2.3) localizes the mouth and both eyes correctly in 100 % images of our database. Table 1 presents the results of this method obtained with r and with radial weights and also results obtained with a relaxed version, not using the limit on the distance of both eyes. These results support the necessity of such condition. Further Table 1 contains results obtained for images with a modified size or rotation, which will be described in Chapter 3.

3.3 Robust correlation

We use the database with 212 images and the 7 mouth templates and 6 eye templates together with their mirror reflections. Several of the robust correlation methods of Chapter 2.4 yield 100 % correct results in the localization of both eyes and the mouth jointly. These correct results are obtained with the LTS-based correlation coefficient with $h = 0.8n$, the LWS-based correlation coefficient with linear weights and also the LWS-based correlation coefficient with adaptive weights. Typically the outliers trimmed by the LTS or down-weighted by the LWS are located in the neighborhood of eyes, at the boundary of the rectangular templates.

For comparison we consider a robust correlation coefficient based on trimming principal variables $u = x + y$ and $v = x - y$, proposed by [27]. This fails in our application in locating the mouth. The reason can be explained on the correlation coefficient between the mouth template and the mouth. The outliers in the variable u are namely typically in the lips, while the outliers in the variable v correspond to the cheeks. For the correlation coefficient between the mouth template and the non-mouth the outliers are typically present in the cheeks. There is no clear correspondence between the outliers in the principal variables and the original data, which does not allow the mouth to be classified correctly.

To summarize, we apply robust versions of the correlation coefficient based on robust regression with a high breakdown point. These give 100 % correct results in a joint localization of both eyes and the mouth, similarly with the weighted correlation coefficient with radial weights.

3.4 Robustness of the methods

Tab. 2. Performance of the weighted Pearson product-moment correlation coefficient with radial weights in locating both eyes and the mouth jointly. Percentages of correct results.

The method is applied on original images, images rotated by any degree and images slightly modified by noise, occlusion or asymmetry as described in Chapter 3.

Images	r_w with radial weights
Original	1.00
Rotated by any degree	1.00
Noise	1.00
Occlusion	1.00
Asymmetry of the face	1.00

We present validation steps to verify the performance of our methods for the joint localization of the mouth and both eyes. We verify the method (4) on a validation set, further we inspect the properties under rotation by any degree and finally we study the robustness empirically by modifying the original images by introducing additional noise, small occlusion or small asymmetry; Table 2 summarizes our modifications of the original images.

This method does not use any parameters learned over the database of images. Specific properties of this particular database influence the choice of the templates and the expected distances between both eyes and between the mouth and eyes. Nevertheless we verify the performance of the method on a validation set of images. We have photographed 30 randomly selected students at the University of Duisburg-Essen with a compact digital camera. Our conditions were standardized to obtain images of faces with the same illumination, distance from the camera, possibly rotated in the plane by small angles, without facial expressions. We transformed the color images from their size 2048*1536 pixels to grey scale images of size 266*200 pixels so that the size of the faces corresponds to the size of the faces in the original database. We used the method of Chapter 2.3 to localize the mouth and both eyes in the images. The method gives 100 % correct results for these images with our set of 7 mouth templates and 6 eye templates. Nevertheless we consider our method sensitive to the size of the images and therefore we do not verify the methods on other databases of images.

Theoretical properties of template matching concerning robustness with respect to rotation, occlusion or asymmetry are studied by [3]. It follows immediately that (4) is robust in the same situations. On the other hand the sample influence function [28] is not limited and (4) turns out to be vulnerable to highly influential weights.

We start by examining the method for the rotated images, while we retain non-rotated templates. Localizing the eyes separately or the mouth alone does not yield successful results in images rotated by $\pm 10^\circ$; not even are the eyes or mouth localized correctly in non-rotated images. Further we compute the joint localization of both eyes and the mouth for images rotated up to $\pm 10^\circ$. Using 6 eye templates, their reflections in the axial symmetry and 7 mouth templates together with radial weights for all mouth and eye templates, the mouth and both eyes are correctly localized in 100 % of the images.

Nevertheless rotating the image by $\pm 20^\circ$ makes the method collapse in about 50 % of the images.

If the face is rotated by any of angles 10, 20, ..., 350 degrees, the sum of three weighted correlation coefficients is smaller than in a non-rotated face with horizontal eyes. A good strategy is to rotate the face by several different angles and to find the largest value of (4) for each rotation. In this way we localize the mouth and both eyes correctly independently on the initial rotation of the given image. To be specific, we rotated each images by angles 0, 10, 20, ..., 180, ..., 350 degrees. The largest value of r_w^* in each of the 212 images is obtained exactly for such position, when the eyes are in a horizontal position. That leads to the correct localization of the mouth and both eyes in 100 % of the images. The 100 % correct results for images rotated by any degree are obtained also for the LWS-based correlation coefficient with linearly decreasing weights and with the adaptive weights.

The method for localizing the mouth and both eyes contains a double protection against a possible rotation of the face. Firstly the templates are robust to a rotation up to $\pm 10^\circ$. Secondly the coefficient r_w^* attains the largest value exactly for a non-rotated face. Therefore the correct rotation of the face is automatically detected in each face with an arbitrary rotation up to $\pm 180^\circ$ degrees. Table 1 presents results obtained with r and r_w with radial weights for images rotated by $\pm 10^\circ$ or reduced in size by 10 %.

Finally we inspect robustness properties to other non-standard situations. Pixel-independent noise with Gaussian distribution with zero expectation does not harm the results with variance up to $\sigma = (0.11)^2$. The mouth and both eyes are correctly localized also in images rotated by any angle (Figure 3) and spoiled by the noise with zero expectation and variance σ^2 up to 0.01.

We occluded the mouth in every image by a small plaster to examine the local sensitivity of the method. Grey values in a rectangle of size 3*5 pixels are set to 1. Every mouth in the database is modified in



Fig. 3. Localization of the mouth and both eyes. For a particular candidate mouth the eyes are searched for in a relevant candidate area.

this way placing the plaster always on the same position to the bottom right corner of the mouth, below the midpoint of the mouth by 7 to 9 rows and on the right from the midpoint by 16 to 20 columns. An example of such occluded mouth is shown in Figure 5. The mouth and both eyes are localized in 100 % of images with such occluded mouths.

To study the effect of asymmetry of the mouth, we increase grey values in the right half of every mouth by a constant, say ϵ . There is no monotone relationship between the value of ϵ and the value of the separation.

The joint search for the mouth and both eyes gives 100 % correct results with ϵ up to 0.15, which is already quite a severe alteration of the original mouth. Here equal and radial weights fail for locating the mouth itself by 7 symmetric mouth templates.

To summarize, the joint localization of mouth and eyes with 7 mouth templates and 6 eye templates has robust properties with respect to small rotation, occlusion and non-symmetry of the image. Robust modifications of the correlation coefficient are applied to the same problem in the next chapter.



Fig. 4. The rotated image additionally modified by noise.

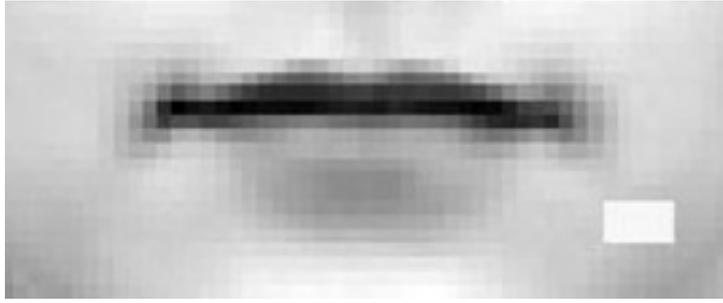


Fig. 5. Mouth modified by a small occlusion (white plaster).

4. Discussion and conclusions

We have proposed a method for the joint localization of the mouth and both eyes in images of faces. The methods are tailor made for the given images of faces with expected usage of the methods in genetic applications. Templates are conceptually simple and clearly interpretable. The method is robust with respect to noise in the images, occlusion or asymmetry of the faces. The results are correct for any initial rotation of the face. The disadvantage of our approach for a general usage is the sensitivity to the size of the images.

Our attempt to apply robust correlation measures to the same task of localizing objects in the image yields promising results and we can recommend robust analogies of the correlation coefficient based on the least trimmed squares and least weighted squares regression for practical usage. Nevertheless the computation of these estimators is very tedious also with approximative algorithms.

The initial denoising was motivated by the necessity to obtain a method robust to noise in the images; its effect may be however small, especially when robust correlation measures with a high breakdown point are used to measure the similarity between the image and the template. While theoretical robustness properties of the robust methods remain for future research, the methods described in this paper can be recommended for practical usage in genetic [17], [18] or anthropologic research [3].

In contrary to standard approaches we examined correlation analysis methods for locating objects in raw images, without a

prior reduction of dimension and feature extraction. Standard approaches use the normalized position or invariant descriptors in order to remove or diminish the effect of rotation or even the change of scale of the images. Nevertheless we contradict a popular belief that statistical methods cannot handle the task of analyzing raw images. Promising results are obtained with the weighted Pearson product-moment correlation coefficient and also robust versions of the Pearson product-moment correlation coefficient. There remains an open problem of computing optimal weights for the weighted correlation coefficient. These should increase the discrimination between those parts of the image corresponding to the template and those which do not. Such problem must be solved again in a robust way, not allowing to obtain highly influential weights by regularizing the problem by introducing a certain upper bound for the optimal weights. An interesting task would be to compare standard classification methods of multivariate statistical analysis and their performance in locating landmarks in the given database of images.

Acknowledgements

This work is supported by Center of Biomedical Informatics, project 1M06014 of the Ministry of Education, Youth and Sports of the Czech Republic. The author is thankful to two anonymous referees for valuable comments and tips for improving the paper.

References

- [1] Yang M.-H., Kriegman D.J., Ahuja N. (2002): Detecting faces in images: A survey. *IEEE Trans. Pattern Anal. and Machine Intel.* 24, No. 1, 34-58.

- [2] Viola P., Jones M.J. (2004): Robust real-time face detection. *Int. Journal of Computer Vision* 57, 137-154.
- [3] Kalina J. (2010): Locating landmarks using templates. *Nonparametrics and Robustness in a Broader Perspectives. A Festschrift in Honor of Professor Jana Jurečková. IMS Collections No. 7.* Accepted, in print.
- [4] Dobeš M., Machala L., Tichavský P., Pospíšil J. (2004): Human eye iris recognition using the mutual information, *Optik* 115, 399-404.
- [5] Lin Z., Davis L.S., Doermann D.S., DeMenthon D. (2007): Hierarchical part-template matching for human detection and segmentation, in *Proceedings of the Eleventh IEEE International Conference on Computer Vision ICCV 2007*, IEEE Computer Society, Washington, 2007, 1-8.
- [6] Wolf L., Huang X., Martin I., Metaxas D. (2006): Patch-Based Texture Edges and Segmentation. In *Leonardis A., Bischof H., Pinz A. (Eds.): Computer Vision - ECCV 2006*, 9th European Conference on Computer Vision, Graz, *Proceedings, Part II. Lecture Notes in Computer Science 3952.*
- [7] Papageorgiou C.P., Oren M., Poggio T. (1998): A general framework for object detection, in *Proceedings of the Sixth IEEE International Conference on Computer Vision ICCV 1998*, IEEE Computer Society, Washington, 555-562.
- [8] Dalal N., Triggs B. (2005): Histograms of oriented gradients for human detection, in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR 2005*, IEEE Computer Society, Washington, 2005, pp. 886-893.
- [9] Torralba A., Murphy K.P., Freeman W.T. (2007): Sharing visual features for multiclass and multiview object detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 5, 854-869.
- [10] Tuzel O., Porikli F., Meer P. (2007): Human detection via classification on Riemannian manifolds, in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR 2007*, IEEE Computer Society, Washington.
- [11] Böhringer S., Vollmar T., Tasse C., Würtz R.P., Gillessen-Kaesbach G., Horsthemke B., Wiczorek D. (2006): Syndrome identification based on 2D analysis software. *Eur. J. Hum. Genet.* 14, 1082-1089.
- [12] Loos H.S., Wiczorek D., Würtz R.P., Malsburg von der C., Horsthemke B. (2003): Computer-based recognition of dysmorphic faces. *Eur. J. Hum. Genet.* 11, 555-560.
- [13] Würtz R.P. (1997): Object recognition robust under translations, deformations, and changes in background. *IEEE Trans. Pattern Anal. and Machine Intel.* 19, No. 7, 769-775.

- [14] Pitas I., Venetsanopoulos A.N. (1990): Nonlinear digital filters. Kluwer, Dordrecht.
- [15] Jurečková J., Pícek J. (2006): Robust statistical methods with R. Chapman & Hall/CRC, Boca Raton.
- [16] Arya K.V., Gupta P., Kalra P.K., Mitra P. (2007): Image registration using robust M-estimators, *Pattern Recognition Letters* 28, 1957-1968.
- [17] Dunning M.J., Smith M.L., Ritchie M.E., Tavaré S. (2007): beadarray: R classes and methods for Illumina bead-based data, *Bioinformatics* 23, 2183-2184.
- [18] Kalina J. (2010): Robust image analysis in the evaluation of gene expression studies. *ERCIM News, European Research Consortium for Informatics and Mathematics*, No. 82, p. 52.
- [19] Víšek J.Á. (2001): Regression with high breakdown point. In Antoch J., Dohnal G. (Eds.): *Proceedings of ROBUST 2000, Summer School of JČMF, JČMF and Czech statistical society*, 324-356.
- [20] Kalina J. (2007): Locating the mouth using weighted templates. *Journal of applied mathematics, statistics and informatics* 3, No. 1, 111-125.
- [21] Čížek P. (2008): Efficient robust estimation of time-series regression models. *Appl. Math.* 53, No. 3, 267-279.
- [22] Rousseeuw P.J., van Driessen K. (1999): A fast algorithm for the minimum covariance determinant estimator, *Technometrics* 41, 212-223.
- [23] Rousseeuw P.J., Leroy A.M. (1987): Robust regression and outlier detection. Wiley, New York.
- [24] Graf H.P., Cosatto E., Gibbon D., Kocheisen M., Petajan E. (1996): Multi-modal system for locating heads and faces. *Second IEEE International Conference on Automatic Face and Gesture Recognition FG 1996*, 88-93.
- [25] James M. (1987): *Pattern recognition*. BSP Professional books, Oxford.
- [26] Seshadri K., Savvides M. (2009): Robust modified active shape model for automatic facial landmark annotation of frontal faces, *Proceedings of the 3rd IEEE international conference on Biometrics: Theory, applications and systems*, IEEE Press, Piscataway, 319-326.
- [27] Shevlyakov G.L., Vilchevski N.O. (2001): *Robustness in data analysis: criteria and methods*. VSP, Utrecht.
- [28] Hampel F.R., Ronchetti E.M., Rousseeuw P.J., Stahel W.A. (1986): *Robust statistics, The approach based on influence functions*. Wiley, New York.

Contact

RNDr. Jan Kalina, Ph.D.

Institute of Computer Science AS CR

Centre of Biomedical Informatics

Pod Vodárenskou věží 2

182 07 Prague 8

Czech Republic

e-mail: kalina@euromise.cz

Reliability of Composite Dichotomous Measurements

Patřicia Martinková¹, Karel Zvára²

¹Centre of Biomedical Informatics, Institute of Computer Science AS CR, Prague, Czech Republic,

²Department of Medical Informatics, Institute of Computer Science AS CR, Prague, Czech Republic

Summary

Reliability of measurement is a measure of its reproducibility under replicate conditions. The classical concept of reliability assumes that measurement Y is composed out of true value T and error term ϵ , two independent random variables, $Y = T + \epsilon$. Reliability of measurement is defined as the ratio of the variance of the true scores to the variance of the observed scores. However, this concept is not applicable in models for dichotomous measurements which do not consider error terms and are instead defined via conditional probabilities. In this paper we examine a more general definition of reliability proposed in [1], which is based on decomposition of variance in mixed effects model. Proposed definition covers the classical definition of reliability and it is, moreover, appropriate for dichotomous measurements, too. Newly, for the proposed definition assumptions are derived, under which the reliability of composite measurement can be predicted by reliability of single measurement (Spearman-Brown formula) and approximate validity of Spearman-Brown formula is shown for the Rasch model. Finally, as a modification of the classical estimate of reliability based on Cronbach's alpha, we examine its counterpart *logistic alpha* introduced in [2], which appears to be more appropriate for composite dichotomous measurements in some cases. Simulations show that the new estimate does not tend to underestimate reliability as often as the Cronbach's alpha does. The new estimate is used in binary data of computerized process of myocardial perfusion diagnosis from cardiac single proton emission computed tomography (SPECT).

Keywords: reliability, binary data, logistic regression, Cronbach alpha, Rasch model, myocardial perfusion diagnosis

1. Introduction and statistical background

Reliability of measurement is a measure of its reproducibility under replicate conditions. In medical practice, the reliability of measurement remains an important topic engendering much discussion. For continuous measurements, reliability analysis and equivalence test for agreement were lately studied by Yi, Wang and He [3]. A nonparametric, probabilistic estimate of reliability used on cognitive tests in Alzheimer's disease was examined by van Belle and Arnold [4]. IRT model-based reliability estimates, which are appropriate for dichotomous or ordinal outcomes, were used in Teresi et. al [5].

Reproducibility studies for binary outcomes are typically analysed using kappa statistics, which was motivated by its relation to the intraclass correlation coefficient [6], [7]. In this paper we take quite a different approach – we discuss the decomposition of variance in mixed effects model settings, as appeared in [8] and propose a new definition of reliability, which covers the classical testing situation and is moreover suitable also for binary data. We also discuss a new estimate of reliability. For better understanding of parallel with classical test theory, a summary of the basic principles of the classical test theory (CTT) is given in this section.

1.1 Reliability of measurement within CTT

In the classical test theory [9], it is assumed that the measurement Y is composed out of the true value T and the error term ϵ , independent continuous random variables

$$\begin{aligned} Y &= T + \epsilon, \\ T &\sim (\mu, \sigma_T^2), \sigma_T^2 > 0, \\ \epsilon &\sim (0, \sigma^2), \sigma^2 > 0. \end{aligned} \quad (1)$$

The reliability of measurement is defined as a ratio of variance of the true score and variance of the observed score

$$reli(Y) = \frac{\text{var}(T)}{\text{var}(Y)} = \frac{\sigma_T^2}{\sigma_T^2 + \sigma^2} = \rho_1. \quad (2)$$

Alternatively, reliability can be defined as a squared correlation between the measured value and the measurement

$$\text{corr}^2(Y, T) = \text{corr}^2(T + \epsilon, T) = \frac{(\sigma_T^2)^2}{(\sigma_T^2 + \sigma^2)\sigma_T^2} = \text{reli}(Y). \quad (3)$$

Also, reliability can be expressed as a correlation between repeated measurements Y_1, Y_2 , that is between two independent and equally accurate measurements of the same true value T

$$Y_j = T + \epsilon_j, \quad j = 1, 2,$$

$$\epsilon_1, \epsilon_2 \sim (0, \sigma^2) \text{ independent}$$

$$\begin{aligned} \text{corr}(Y_1, Y_2) &= \text{corr}(T + \epsilon_1, T + \epsilon_2) = \\ &= \frac{\sigma_T^2}{\sigma_T^2 + \sigma^2} = \text{reli}(Y) \end{aligned} \quad (4)$$

Correlation between two independent (not necessary equally accurate) measurements of the same true value can be expressed as

$$Y_j = T + \epsilon_j, \quad j = 1, 2,$$

$$\epsilon_1 \sim (0, \sigma_1^2), \epsilon_2 \sim (0, \sigma_2^2) \text{ independent}$$

$$\begin{aligned} \text{corr}(Y_1, Y_2) &= \text{corr}(T + \epsilon_1, T + \epsilon_2) = \\ &= \frac{\sigma_T^2}{\sqrt{\sigma_T^2 + \sigma_1^2} \sqrt{\sigma_T^2 + \sigma_2^2}} = \sqrt{\text{reli}(Y_1)} \sqrt{\text{reli}(Y_2)} \end{aligned} \quad (5)$$

1.2 Reliability of sum of repeated measurements

Having J repeated measurements of T

$$Y_j = T + \epsilon_j, \quad \epsilon_j \sim (0, \sigma^2) \text{ iid}, \quad j = 1, \dots, J, \quad (6)$$

variability of their sum $Y_* = \sum_j Y_j$ is

$$\text{var}(Y_{\bullet}) = \text{var}\left(\sum_j (T + \epsilon_j)\right) = J^2\sigma_T^2 + J\sigma^2, \tag{7}$$

hence reliability ρ_j of sum Y_{\bullet} , or of average

$$\bar{Y}_{\bullet} = \frac{1}{J}Y_{\bullet}$$

can be expressed by means of reliability of single measurement ρ_1 (Spearman-Brown formula, see [10], [11]).

$$\begin{aligned} \text{reli}(Y_{\bullet}) &= \frac{\text{var}(JT)}{\text{var}(Y_{\bullet})} = \frac{J^2\sigma_T^2}{J^2\sigma_T^2 + J\sigma^2} = \\ &= \frac{J\rho_1}{(J-1)\rho_1 + 1} = \rho_J. \end{aligned} \tag{8}$$

Reliability of sum of J repeated measurements can also be expressed in a more practical way

$$\begin{aligned} \text{reli}(Y_{\bullet}) &= \frac{\text{var}(JT)}{\text{var}(Y_{\bullet})} = \frac{J}{J-1} \frac{J-1}{J} \frac{J^2\sigma_T^2}{\text{var}(Y_{\bullet})} = \\ &= \frac{J}{J-1} \frac{J^2\sigma_T^2 + J\sigma^2 - J\sigma^2 - J\sigma_T^2}{\text{var}(Y_{\bullet})} \\ &= \frac{J}{J-1} \frac{\text{var}\left(\sum_j Y_j\right) - \sum_j \text{var}(Y_j)}{\text{var}\left(\sum_j Y_j\right)} = \\ &= \frac{J}{J-1} \frac{\sum_{j \neq k} \sigma_{jk}}{\sum_{j,k} \sigma_{jk}} = \alpha \end{aligned} \tag{9}$$

where $\sigma_{jk} = \text{cov}(Y_j, Y_k)$. In equation (9), we got so called **Cronbach's alpha** (see [12]) which may be estimated by using sample covariances instead of their population counterparts

$$\hat{\alpha} = \frac{J}{J-1} \frac{\sum_{j \neq k} s_{jk}}{\sum_{j,k} s_{jk}},$$

$$\text{kde } s_{jk} = \frac{1}{I-1} \sum_{t=1}^I (Y_{tj} - \bar{Y}_{\bullet j})(Y_{tk} - \bar{Y}_{\bullet k}). \tag{10}$$

For dichotomous data, estimate (10) coincides with Kuder-Richardson formula 20 (see [13]).

As shown above in (9), Cronbach's alpha is

equivalent to reliability of sum of repeated measurements of the *same* true value T . Nevertheless, it is also widely used as an estimator of reliability of composite measurements.

1.3 Reliability of composite measurements

Often, the measurement cannot be repeated independently to produce exactly the same true value T . In psychometrics, the tests are composed of J items where each of concentrates on a slightly different aspect of measured quantity T_j ; each subject is described by the sum of J item scores $Y_{\bullet} = \sum Y_j$.

In medical practice, the health professionals are often faced with the same quantitative measurements reported by different raters, or from the same rater measured using different tools, and the measured property is often described by the average of the measurements

$$\bar{Y}_{\bullet} = \sum Y_j / J$$

For j -th measurement, we suppose that

$$Y_j = T_j + \epsilon_j, \quad \epsilon_j \sim (0, \sigma_j), \tag{11}$$

$$j = 1, \dots, J$$

where $(\epsilon_1, \dots, \epsilon_J)$ are mutually independent and also independent of (T_1, \dots, T_J)

Reliability of composite measurement

$$Y_{\bullet} = \sum_j Y_j$$

is

$$\text{reli}(Y_{\bullet}) = \frac{\text{var}(T_{\bullet})}{\text{var}(T_{\bullet}) + \text{var}(\epsilon_{\bullet})}. \tag{12}$$

As demonstrated by Novick and Lewis in [14], Cronbach's alpha is generally a lower bound of reliability of composite measurement

$$\alpha \leq \text{reli}(Y_{\bullet}) = \rho_J. \tag{13}$$

Novick and Lewis showed that equality holds only for essentially τ -equivalent items, that is in case, where for a random variable T and real numbers β_j such that

$$\sum \beta_j = 0,$$

with probability = 1 holds

$$T_j = T + \beta_j, \quad \forall j = 1, \dots, J. \tag{14}$$

This is equivalent to simultaneously holding

$$\text{var}(T_j) = \sigma_T^2, \quad \forall j \tag{15}$$

$$\text{corr}(T_j, T_k) = 1, \quad \forall j, k. \tag{16}$$

Condition (1.3) is needed to have equal item reliabilities in model (11). For Spearman-Brown formula to hold, also condition (16) is needed.

When items of composite measurement are not essentially τ -equivalent, besides the fact, that Spearman-Brown formula does not hold, from (13) we might also expect that estimate of reliability based on Cronbach's alpha will underestimate the true reliability. Some estimations of this discrepancy between reliability of composite measurement and Cronbach's alpha on population level can be found in [15].

1.4 Cronbach's alpha as estimator of reliability of composite measurements

Let us now suppose measurements Y_{ij} , $j = 1, \dots, J$ on subjects $i = 1, \dots, I$. Assumptions of essential τ -equivalence lead to 2-way mixed effects ANOVA model.

$$Y_{ij} = T_i + \beta_j + \epsilon_{ij}, \quad \sum_j \beta_j = 0, \tag{17}$$

$$i = 1, \dots, I, \quad j = 1, \dots, J$$

where β_j is the item parameter (item difficulty, expert's level, etc.). When we add assumptions of normality

$$T_i \sim N(\mu, \sigma_T^2), \quad \epsilon_{ij} \sim N(0, \sigma^2),$$

and consider sum of squares decomposition

$$SS_T = \sum (Y_{ij} - \bar{Y}_{\bullet\bullet})^2 = SS_A + SS_B + SS_E,$$

the mean squares MS_A and MS_E have the following expectations

$$EMS_A = E \sum_{i=1}^I \sum_{j=1}^J (\bar{Y}_{i\bullet} - Y_{i\bullet})^2 / (I - 1) = J\sigma_T^2 + \sigma^2, \tag{18}$$

$$EMS_E = E \sum_{i=1}^I \sum_{j=1}^J (Y_{ij} - \bar{Y}_{i\bullet} - \bar{Y}_{\bullet j} + \bar{Y}_{\bullet\bullet})^2 / ((I - 1)(J - 1)) = \sigma^2 \tag{19}$$

Hence, Cronbach's alpha (9) can be expressed as

$$\alpha = \frac{EMS_A - EMS_E}{EMS_A} \tag{20}$$

and the estimate (10) can be rewritten as

$$\hat{\alpha} = \frac{MS_A - MS_E}{MS_A} = 1 - \frac{MS_E}{MS_A} = 1 - \frac{1}{F_T}, \tag{21}$$

where F_T is the statistic used for submodel testing with no subject effect in the full model (17). The same statistic is used also for testing the submodel in 2-way ANOVA fixed effects model [16].

From (21) we can conclude that high values of $\hat{\alpha}$ indicate that the composite measurement can distinguish between the subjects well. Hence, we may expect higher values of $\hat{\alpha}$ when the subjects' variability is high.

2. New developments for dichotomous measurements

When measurements Y takes only the values 0 or 1, the classical model described in part 1.1 is not appropriate anymore, since Y cannot be expressed as a sum of two independent random variables. Instead, the model should be defined through conditional mean values $E(Y|T)$. Let us suppose a more general model

$$Y_{ij} \sim f(\cdot, T_i), \quad T_i \sim iid,$$

$$(Y_{ij}|T_i), (Y_{i'j'}|T_{i'}) \text{ independent for } j \neq j' \tag{22}$$

One of such models is the Rasch model (see [17]):

$$E(Y_{ij}|T_i) = \pi(T_i, \beta_j) = \frac{\exp(T_i + \beta_j)}{1 + \exp(T_i + \beta_j)},$$

$$T_i \sim N(\mu, \sigma_T^2). \tag{23}$$

In the framework of Item Response Theory (IRT) the Rasch model (23) and its generalizations are widely studied [18], especially in connection with parameter estimation. The concept of reliability is extended from a single index to a function of the true value T called the *test information function* [19]. Besides, it is also possible to obtain an index for a test as a whole which is directly analogous to Cronbach's alpha: As an analogy to decomposing an observed score into true score and an error in classical test theory, we consider a decomposition of an IRT person estimate into a true location and error

$$\hat{T} = T + \epsilon. \tag{24}$$

The reliability in IRT is defined as

$$R_{IRT} = \frac{\text{var}(T)}{\text{var}(\hat{T})} = \frac{\text{var}(\hat{T}) - \text{var}(\epsilon)}{\text{var}(\hat{T})}. \tag{25}$$

For its estimation, first, the estimates of subjects' true locations and their standard errors (SE) are computed by standard estimation procedures. Then, the sample variance of these estimates is computed to estimate $\text{var}(\hat{T})$. The mean squared subject standard error estimate provides an estimate of the variance of the error $\text{var}(\epsilon)$. The resulting estimate is typically very close to Cronbach's alpha [20].

In this paper, we consider a different definition of reliability for model (22), one that is a more straightforward generalization of classical reliability. We give formula for reliability in the Rasch model and assumptions for Spearman-Brown formula to hold in model (22). We also introduce a new estimate of reliability appropriate for composite dichotomous measurements and we compare it to the Cronbach's alpha in simulations and in practical example.

2.1 Proposed definition of reliability

The total observed variance $\text{var}(Y_{ij})$ can be decomposed by the means of conditional variance and conditional mean value as

$$\text{var}(Y_{ij}) = E(\text{var}(Y_{ij}|T_i)) + \text{var}(E(Y_{ij}|T_i)). \tag{26}$$

where the first term is the intraclass variance, that is the part of the variance, which is not due to the variability of T_i and the second term is the interclass variance, the part of total variance which is due to the variability of T_i [8].

To follow the definition of reliability from the classical test theory, we might define it as a ratio of the variance due to variability of the measured property T to the total observed variability, that is

$$\text{reli}(Y) = \frac{\text{var}(E(Y|T))}{\text{var}(Y)} = \frac{\text{var}(E(Y|T))}{\text{var}(E(Y|T)) + E(\text{var}(Y|T))}. \tag{27}$$

Since for the classical model holds

$$E(Y|T) = E((T + E)|T) = T,$$

definition (27) coincides with the classical definition (2).

For composite measurements, the reliability of j -th item can be defined as

$$\text{reli}(Y_{ij}) = \frac{\text{var}(E(Y_{ij}|T_i))}{\text{var}(Y_{ij})} = \tau_{ij} \tag{28}$$

and the reliability of composite measurement $Y_{i\bullet}$ can be defined as

$$\text{reli}(Y_{i\bullet}) = \frac{\text{var}(E(Y_{i\bullet}|T_i))}{\text{var}(Y_{i\bullet})}. \tag{29}$$

For the Rasch model (23), we derived in [1] that reliability of composite measurement is

$$\text{reli}(Y_{i\bullet}) = \frac{\text{var}(E(Y_{i\bullet}|T_i))}{\text{var}(Y_{i\bullet})} = \frac{\sum_{j=1}^J \sum_{t=1}^J (C_{jt} - D_j D_t)}{\sum_{j=1}^J \sum_{t=1}^J (C_{jt} - D_j D_t) + \sum_{j=1}^J B_j}, \tag{30}$$

where

$$B_j = E_T \frac{e^{T+\beta_j}}{(1 + e^{T+\beta_j})^2},$$

$$D_j = E_T \frac{e^{T+\beta_j}}{1 + e^{T+\beta_j}},$$

$$C_{j'l} = E_T \frac{e^{T+\beta_j}}{1 + e^{T+\beta_j}} \frac{e^{T+\beta_l}}{1 + e^{T+\beta_l}}$$

These integrals cannot be evaluated explicitly, nevertheless they can be evaluated numerically. Hence, for a given testing situation (that is for distribution of subjects' true values T number of items J and their levels $\beta_j, j = 1, \dots, J$) the true value of reliability can be computed (see Table 1).

2.2 Spearman-Brown formula

Preliminary let's find assumptions for having equal item reliabilities τ_j for all j . As we have already mentioned, reliability can also be expressed as a correlation of two independent measurements of the same property T (see formula (4)), that is by the (i -th subject) intraclass correlation

$$\rho_{ijj'} = \text{corr}(Y_{ij}Y_{ij'}).$$

Within model (22), the relationship between $\rho_{ijj'}$ and $\tau_j, \tau_{j'}$ for $j \neq j'$ is following (for proof, see [21]):

$$\begin{aligned} \rho_{ijj'} &= \sqrt{\tau_{ij}} \sqrt{\tau_{ij'}} \text{corr}[E(Y_{ij}|T_i), \\ E(Y_{ij'}|T_i)] &\leq \sqrt{\tau_{ij}} \sqrt{\tau_{ij'}} \end{aligned} \quad (31)$$

The equality in (31) holds for all $j \neq j'$ if for all $j \neq j'$

$$\text{corr}[E(Y_{ij}|T_i), E(Y_{ij'}|T_i)] = 1, \quad (32)$$

that is, in the case when for all i, j with probability equal to one, for some constants $k_j > 0$ and η_j and some functions $\lambda_i(T)$ the conditional means can be expressed as

$$E(Y_{ij}|T_i) = k_{ij} [\lambda_i(T_i) + \eta_{ij}]. \quad (33)$$

Moreover, assumption (33) can be required with additional constraint

$$\sum_{k=1}^J k_{ij}^2 = J,$$

since λ_i and η_j can be multiplied by appropriate constants.

Formula (31) and assumption (33) may be extended to the following theorem (for proof, see [21]), which revises the theorem of Commenges and Jacqmin (see [8]):

Theorem Suppose that Y_{ij} for $i = 1, \dots, I, j = 1, \dots, J, J \geq 3$ obey the model (22). Moreover, with probability equal to one let for all i, j hold the assumption (33). Then the following propositions are equivalent:

P1 $\rho_{ijj'} = \rho_i$ does not depend on j, j' for any $j \neq j'$.

P2 $\tau_j = \tau_i$ does not depend on j for any j, j' .

P3 The model belongs to a class specified (with probability equal to one) by:

$$\text{var}(Y_{ij}|T_i) = k_{ij}^2 [\sigma_i^2(T_i) + \psi_{ij}(T_i)], \quad (34)$$

where $E[\psi_{ij}(T_i)] = 0$, and $\psi_{ij}(T_i) > -\sigma_i^2(T_i)$.

P4 $\rho_i = \tau_i$.

Hence, when assumptions (33) and (34) hold, the item reliabilities equal τ_i for all j . Let us now look at reliability of composite measurements under these two assumptions and under mentioned constraint

$$\sum_{j=1}^J k_{ij}^2 = J.$$

The reliability of (every) single item may be written as

$$\begin{aligned} R_1 = \tau_{ij} &= \frac{\text{var}[E(Y_{ij}|T_i)]}{\text{var}[E(Y_{ij}|T_i)] + E[\text{var}(Y_{ij}|T_i)]} = \\ &= \frac{\text{var}[\lambda_i(T_i)]}{\text{var}[\lambda_i(T_i)] + E[\sigma_i^2(T_i)]} = \tau_i = \rho \end{aligned}$$

and by Theorem it coincides with the correlation of two independent measurements of the same property $Y_{ij}, Y_{ij'}, j \neq j'$. The reliability of the composite measurement is

$$\begin{aligned} R_J &= \frac{\text{var}[E(Y_i|T_i)]}{\text{var}[E(Y_i|T_i)] + E[\text{var}(Y_i|T_i)]} = \\ &= \frac{\text{var}[\sum_{j=1}^J E(Y_{ij}|T_i)]}{\text{var}[\sum_{j=1}^J E(Y_{ij}|T_i)] + \sum_{j=1}^J E[\text{var}(Y_{ij}|T_i)]} \\ &= \frac{\text{var}\{\sum_{j=1}^J k_{ij}[\lambda_i(T_i) + \eta_{ij}]\}}{\text{var}\{\sum_{j=1}^J k_{ij}[\lambda_i(T_i) + \eta_{ij}]\} + \sum_{j=1}^J k_{ij}^2 E[\sigma_i^2(T_i) + \psi_{ij}(T_i)]} \\ &= \frac{(\sum_{j=1}^J k_{ij})^2 \text{var}[\lambda_i(T_i)]}{(\sum_{j=1}^J k_{ij})^2 \text{var}[\lambda_i(T_i)] + m E[\sigma_i^2(T_i)]} = \\ &= \frac{(\sum_{j=1}^J k_{ij})^2 R_1}{1 + \left(\frac{\sum_{j=1}^J k_{ij}}{J} - 1\right) R_1}. \end{aligned} \quad (35)$$

The expression (35) coincides with Spearman-Brown formula if $k_j = 1$ for all j . We may conclude that the assumptions (15)(16) of essential τ -equivalence for classical model correspond in model (22) with assumption that with probability equal to one, the conditional mean and variance of Y_{ij} may be written as

$$E(Y_{ij}|T_i) = \lambda_i(T_i) + \eta_{ij}, \quad (36)$$

$$\text{var}(Y_{ij}|T_i) = \sigma_i^2(T_i) + \psi_{ij}(T_i), \quad (37)$$

where η_j are given constants,

$$E[\psi_{ij}(T_i)] = 0,$$

and $\psi_{ij}(T_i) > -\sigma_i^2(T_i)$.

As showed in [21], the Rasch model (23) does not follow assumption (36) nor (37). Nevertheless, Table 1 and Table 2 give us an impression that the Spearman-Brown formula (8) does hold at least approximately.

In Table 1, the values of reliability were calculated from formula (30) for different testing situations (number of items J item levels equidistant on $\langle -0.1, 0.1 \rangle$ subject levels $T_i \sim N(0, \sigma_T^2)$).

To evaluate the integrals, function integrate in software R (see [22]) was used.

The maximum absolute error reached in integrations was less than 0.000025. In Table 2, we set $J = 11$ and used the second line of Table 1 together with the Spearman-Brown formula (8) to get approximate values of reliabilities for $J = 3, 20, 50$ and 100.

As we may see, the numerical values in Tables 1 and 2 are very similar. As an explanation we give an approximation of Spearman-Brown formula for the Rasch model. Let us assume b_j small,

$$\sum_j b_j = 0,$$

and apply the first-order Taylor series approximation (to function of one or two variables)

$$B_j = E \frac{e^{T+b_j}}{(1+e^{T+b_j})^2} \approx E \frac{e^T}{(1+e^T)^2} +$$

$$+ b_j E \frac{e^T(1-e^T)}{(1+e^T)^3} = B + b_j E \frac{e^T(1-e^T)}{(1+e^T)^3},$$

$$D_j = E \frac{e^{T+b_j}}{1+e^{T+b_j}} \approx E \frac{e^T}{1+e^T} +$$

$$+ b_j E \frac{e^T}{(1+e^T)^2} = D + b_j E \frac{e^T}{(1+e^T)^2},$$

$$C_{jt} = E \frac{e^{T+b_j}}{1+e^{T+b_j}} \frac{e^{T+b_t}}{1+e^{T+b_t}} \approx$$

$$\approx E \frac{e^{2T}}{(1+e^T)^2} + (b_j + b_t) E \frac{e^{2T}}{(1+e^T)^3} =$$

$$= C + (b_j + b_t) E \frac{e^{2T}}{(1+e^T)^3}.$$

Then, the reliability of the composite measurement in the Rasch model is approximately

$$R_J \approx \frac{J^2(C - D^2)}{J^2(C - D^2) + JB}.$$

Tab. 1. Reliability in the Rasch model for different number of items.

Number of items	Variability of subjects σ_T^2						
	0.01	0.1	0.2	0.5	0.9	2.5	10
J=3	0.00008	0.00741	0.02881	0.15047	0.34335	0.73121	0.94152
J=11	0.00028	0.02667	0.09814	0.39386	0.65731	0.90890	0.98335
J=20	0.00050	0.04747	0.16519	0.54160	0.77717	0.94775	0.99077
J=50	0.00125	0.11078	0.33098	0.74709	0.89711	0.97843	0.99629
J=100	0.00249	0.19947	0.49735	0.85524	0.94577	0.98910	0.99814

Tab. 2. Spearman-Brown formula used for $J=11$.

Number of items	Variability of subjects σ_T^2						
	0.01	0.1	0.2	0.5	0.9	2.5	10
SB R_3	0.00008	0.00742	0.02882	0.15054	0.34345	0.73125	0.94153
SB R_{20}	0.00050	0.04746	0.16518	0.54159	0.77716	0.94775	0.99077
SB R_{50}	0.00125	0.11077	0.33095	0.74707	0.89710	0.97843	0.99629
SB R_{100}	0.00249	0.19944	0.49731	0.85522	0.94576	0.98910	0.99814

While the reliability of single measurement is

$$R_1 \approx \frac{C - D^2}{C - D^2 + B},$$

which together gives an approximate validity of the Spearman-Brown formula in the Rasch model.

2.3 Estimate of reliability logistic alpha

F_T statistic in estimate based on Cronbach's alpha (21) is best suited for normally distributed data. For dichotomous data we might think of replacing F_T by analogous statistic from logistic regression. In the fixed effects model of logistic regression, the appropriate statistic is the difference of deviances in the submodel and in the model $X^2 = D(B) - D(A+B)$. This statistic has under the null hypothesis asymptotically (for I fixed and approaching infinity) the χ^2 distribution with χ^2 degrees of freedom. Hence, the proposed estimate of reliability for composite dichotomous measurements, logistic alpha [2], [1] is:

$$\hat{\alpha}_{log} = 1 - \frac{I - 1}{X^2}. \quad (38)$$

In the following section we compare the new estimate logistic alpha to the classical estimate based on Cronbach's alpha.

3. Simulation example

This simulation is dedicated to the following example: $I = 20$ patients answered to $J = 20$ yes/no items of quality of life survey. The item levels β_j were supposed to be equidistant on $\langle -0.1, 0.1 \rangle$, patients' true values of QOL were assumed to be normally distributed with mean $\mu = 0$ and variance σ_T^2 (55 values of σ_T^2 were chosen from interval $\langle 0.01, 10 \rangle$ to get 55 values of reliability approximately uniformly distributed on interval).

For each combination of I, J and the true reliability was enumerated by formula (30) and 500 data sets generated: Set of I patients' life quality levels T_i was generated from $N(0, \sigma_T^2)$. QOL survey answers Y_{ij} were generated from the Rasch model (23) and estimates $\hat{\alpha}_{cR}$ and $\hat{\alpha}_{log}$ were computed from the data.

From obtained 500 estimates $\hat{\alpha}_{cR}$ and 500 estimates $\hat{\alpha}_{log}$ the bias and mean squared error (MSE) were computed and plotted out in Figure 1 and 2.

Other testing situations (number of items $J = 11$ and $I = 30$ and number of patients $I = 30$ and 50) were studied, too. We observed smaller bias and MSE in $\hat{\alpha}_{og}$ particularly for true reliability ≤ 0.75 . Inferior results of the *logistic alpha* were obtained for reliability close to 1 and for high number of patients in proportion to the number of items. The latter might be a consequence of the fact, that while statistic X^2 used in (38) is appropriate for fixed effects model of logistic regression, in (23) we expect a mixed effects model. A chance of improvement lies in replacing statistic X^2 by even more appropriate one.

4. Analysis of cardiac data

The dataset **SPECT heart data** [23] describes diagnosing of cardiac Single Proton Emission Computed Tomography (SPECT) images. Each of the patients is classified into two categories: normal and abnormal. The database of 267 SPECT image sets (patients) was processed to extract features that summarize the original SPECT images. As a result, 22 binary feature patterns (22 partial diagnoses based on SPECT) were created for each patient [24]. We were interested in internal consistency of the 22 partial diagnoses based on SPECT. Reached estimates of reliability were: Cronbach alpha $\hat{\alpha}_{cR} = 0.839$ and logistic alpha $\hat{\alpha}_{og} = 0.827$. In this case, both estimates are quite similar and show on high internal consistency of partial diagnoses.

5. Conclusion and discussion

In this paper, the basic principles from classical test theory (CTT) were summarized and used for a new definition of reliability and new estimate of reliability appropriate for composite dichotomous measurements. For classical testing situation the proposed definition of reliability was shown to coincide with definition of reliability in CTT. Assumptions for Spearman-Brown formula were given for model (22) which is more general than classical model (1). The proposed definition and estimate of reliability were applied in the Rasch model, for which the Spearman-Brown formula was shown to hold only approximately.

The proposed estimate *logistic alpha* was shown to possess better properties (smaller bias and MSE), in particular for true reliability ≤ 0.75 and the number of items exceeding the number of patients. The chances of improvement of the new estimate for true reliability close to 1 and for higher number of patients were discussed. Estimation of reliability on binary data was demonstrated on cardiac data. Work presented in this paper could lead to more precise estimation of reliability for binary data, which could contribute to many fields of biomedical research.

References

- [1] Martinková P, Zvára K. Reliability in the Rasch model. *Kybernetika* 2007; 43(3):315326.
- [2] Zvára K. Measuring of reliability: Beware of Cronbach. [Měření reliability aneb bacha na Cronbacha, in Czech]. *Information Bulletin of the Czech Statistical Society* 2002; 12:1320.
- [3] Yi Q, Wang PP, He Y. Reliability analysis for continuous measurements: Equivalence test for agreement. *Statistics in Medicine* 2008; 27:28162825.
- [4] vanBelle G, Arnold A. Reliability of cognitive tests used in Alzheimer's disease. *Statistics in Medicine* 2000; 19:14111420.
- [5] Teresi JA, Kleinman M, Ocepek-Weliksion K. Modern psychometric methods for detection of differential item functioning: application to cognitive assessment measures. *Statistics*

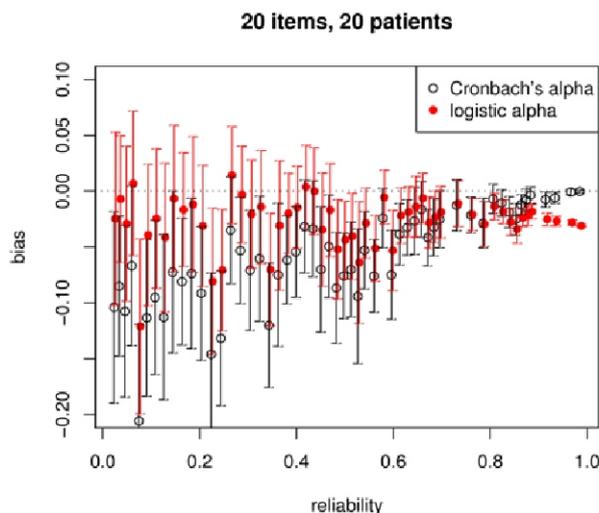


Fig. 1: Estimated Bias and its confidence interval for classical and logistic estimator of reliability

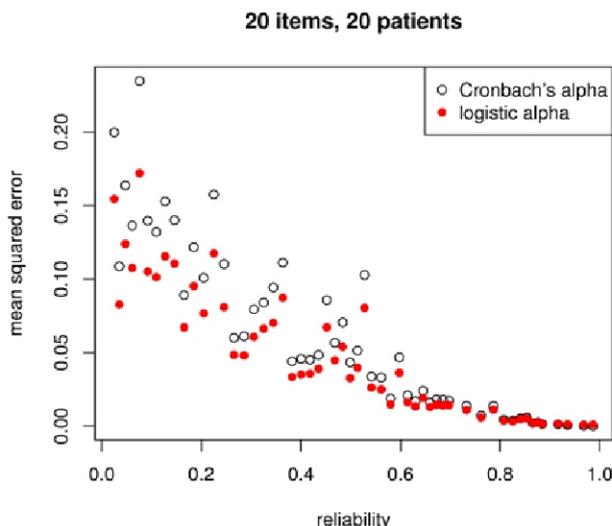


Fig. 2: Estimated MSE for classical and logistic estimator of reliability

- [6] Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educational and Psychological Measurement* 1973; 33: 613619.
- [7] Kraemer HC. Ramifications of a population-model for kappa as a coefficient of reliability. *Psychometrika* 1979; 44(4):461472.
- [8] Commenges D, Jacqmin H. The intraclass correlation coefficient distribution-free definition and test. *Biometrics* 1994; 50:517526.
- [9] Suen HK. *Principles of Test Theories*. LEA Publishers, Hillsdale, New Jersey, 1990.
- [10] Spearman C. Correlation calculated from faulty data. *British Journal of Psychology* 1910; 3:271296.
- [11] Brown W. Some experimental results in the correlation of mental abilities. *British Journal of Psychology* 1910; 3:296322.
- [12] Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951; 16:297334.
- [13] Kuder G, Richardson M. The theory of estimation of test reliability. *Psychometrika* 1937; 2:151160.
- [14] Novick MR, Lewis C. Coefficient alpha and the reliability of composite measurement. *Psychometrika* 1967; 32:113.
- [15] Raykov T. Scale reliability, Cronbach's coefficient Alpha, and violations of essential tau-equivalence with fixed congeneric components. *Multivariate Behavioral Research* 1997; 32(4):329353.
- [16] Neter J, Wasserman W, Kutner MH. *Applied Linear Statistical Models*. Richard D. Irwin, Inc., Homewood, IL, USA, 1985
- [17] Rasch G. *Probabilistic Models for Some Intelligence and Attainment Tests*. The Danish Institute of Educational Research, 1960.
- [18] van der Linden WJ, Hambleton RK (editors). *Handbook of Item Response Theory*. Springer-Verlag, New York, 1997.
- [19] Samejima F. Estimation of Reliability Coefficients Using the Test Information Function and Its Modifications. *Applied Psychological Measurement* 1994; 18(3):229244.
- [20] Andrich D. An index of person separation in latent trait theory, the traditional KR-20 index, and the Guttman scale response pattern. *Educational Research and Perspectives* 1982; 9(1):95104.
- [21] Martinková P. *Reliability of Measurements Consisting of Dichotomously Scored Items*. Unpublished dissertation. Charles University, Prague, 2007.
- [22] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2010. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- [23] Cios KJ, Kurgan LA. *SPECT heart data*. Colorado, USA, 2010. URL <http://archive.ics.uci.edu/ml/datasets/SPECT+Heart>.
- [24] Kurgan LA et al. Knowledge discovery approach to automated cardiac SPECT diagnosis. *Artificial Intelligence in Medicine* 2001; 23: 149-169.

Contact

RNDr. Patřicia Martinková, Ph.D.
Institute of Computer Science AS CR
Centre of Biomedical Informatics
Pod Vodárenskou věží 2
182 07 Prague 8
Czech Republic
e-mail: martinkova@euromise.cz



EJBI 2010 ISSN 1801 - 5603

European Journal for Biomedical Informatics

Volume 6 (2010), Issue 2

Czech version

Cooperating journal – Lékař a technika

Editor-in-Chief: Karel Roubík (Czech Republic)

Publisher: Společnost biomedicínského inženýrství a lékařské informatiky ČLS JEP
a Fakulta biomedicínského inženýrství ČVUT v Praze

www.ejbi.eu



Content

Czech version

- cs 91 [Předmluva](#)
Jana Zvárová
- cs 92 - 94 [Pracovní skupina zaměřená na časopisy Evropské federace pro medicínskou informatiku: pozadí, odůvodnění a účel](#)
Izet Masic, Josipa Kern, Jana Zvárová, Simon de Lusignan
- cs 95 - 102 [Robustní analýza obrazu obličeje pro genetické aplikace](#)
Jan Kalina
- cs 103 - 109 [Spolehlivost složených dichotomních měření](#)
Patricia Martinková, Karel Zvára

Předmluva

Biomedicínská informatika je rychle se rozvíjející obor, s důležitými aplikacemi a dopady na celý biomedicínský svět a poskytování zdravotnické péče. European Journal for Biomedical Informatics (EJBI) reaguje na velkou evropskou potřebu sdílet informace v mnohojazyčné a multikulturní Evropě.

V roce 2010 se EJBI stalo oficiálním časopisem Evropské federace pro medicínskou informatiku (EFMI). Těmito

změnami časopis zdůrazní roli informačních technologií a národních časopisů a podpoří širokou komunikaci v mnoha evropských jazycích.

Role národních a mezinárodních evropských časopisů medicínské informatiky je velice důležitá. V prvním článku druhého čísla EJBI 2010 je detailně popsána role EFMI a evropských časopisů medicínské informatiky. Druhé dva články tohoto čísla se zaměřují na robustní

analýzu obrazu obličeje pro genetické aplikace a na spolehlivost složených dichotomiích měření.

*Jana Zvárová
vedoucí redaktorka*

Pracovní skupina zaměřená na časopisy Evropské federace pro medicínskou informatiku: pozadí, odůvodnění a účel

Izet Masic¹, Josipa Kern², Jana Zvárová³, Simon de Lusignan⁴

¹Society for Medical Informatics of Bosnia and Herzegovina, Sarajevo, Bosnia and Herzegovina,

²Croatian Society for Medical Informatics, Zagreb, Croatia,

³Czech Society for Biomedical Engineering and Medical Informatics, Prague, Czech Republic,

⁴Division of Community Health Sciences, St George's University of London, UK

Přetištěno z: *Medical Informatics in a United and Healthy Europe*, K.P. Adlassnig et al. (Eds), IOS Press, 2009, 946-950, © 2009 European Federation for Medical Informatics.

Souhrn

Příspěvek prezentuje analýzu toho, jak EFMI šíří nové znalosti a analýzu aktivních časopisů medicínské informatiky v členských zemích EFMI. Tato analýza byla výsledkem jednání Rady EFMI v Londýně v roce 2008. Analýza identifikuje osm hlavních aktivních informatických časopisů a několik dalších publikací. Většina z nich je založena na předplatném a vycházejí alespoň jednou za čtvrt roku. Existuje možnost, aby se editoři pravidelně scházeli a vytvořili praktickou společnost s cílem dále zlepšovat jejich účinnost v šíření nových znalostí a osvědčených postupů v oblasti medicínské informatiky. Je možné sdílet odborné znalosti a bylo by možné sladit několik aspektů příprav a podávání rukopisů a tím by byly sníženy některé ze zjištěných překážek v publikování.

1. Úvod

Evropská federace pro medicínskou informatiku (EFMI) je nezisková vědecká a profesionální organizace, založená v roce 1976, kterou v současné době tvoří 31 národních členů [1-4]. Všechny evropské státy jsou oprávněny být v EFMI zastoupeny příslušnou národní společností medicínské informatiky. Termín „*medicínská informatika*“ zahrnuje celé spektrum zdravotnické informatiky a všechny její podobory a příbuzné obory. Medical Subject Headings (MeSH) definuje medicínskou informatiku jako [4], [5]: „*Oblast informatiky zabývající se analýzou a šířením medicínských dat přes počítačové aplikace až po různé aspekty zdravotnictví a medicíny*“. Tato definice byla zavedena v roce 1987 jako MeSH termín.

EFMI pracuje s minimální byrokratickou režii, každá národní společnost podporuje federaci tím, že vyšle svého zástupce, aby se zapojil do Rady. Angličtina byla přijata jako oficiální jazyk, i když pro kongresy v neanglicky mluvících zemích je často poskytován i simultánní překlad [6]. EFMI má 16 pracovních skupin: „Paměťové karty ve zdravotnictví „CARDS“; Sociální zabezpečení a blahobyt; Vzdělávání ve zdravotnické informatice „EDU“; Elektronické zdravotní záznamy „EHR“; Evaluace a hodnocení zdravotnických informačních systémů „EVAL“; Informatika pro postižené osoby a rehabilitaci „IDR“; Plánování a modelování informací ve zdravotnictví „IPAM“; Volně dostupný software „LIFOSS“; Směs případů a vážnost případů; Medicínská informatika v tranzitivních státech „MICIT“; Zpracování medicínských obrazů „MIP“; Porozumění přirozenému jazyku „NLU“; Ošetrovatelská informatika v Evropě „NURSIE“; Lidské a organizační faktory medicínské informatiky „HOFMI“; Informatika primární péče „PCI“; Bezpečnost, zabezpečení a etika „SSE“ a Zpětná zjistitelnost „TRA“.

K dnešnímu dni zorganizovala EFMI 21 evropských kongresů medicínské informatiky, včetně kongresů Medical Informatics Europe (MIE) v Cambridgi (1978), Berlíně (1979), Oslu (1988), Glasgow (1990), Vídni (1991), Jeruzalémě (1993), Lisabonu (1994), Kodani (1996), Thessaloniki (1997), Ljubljani (1999), Honnoveru (2000), Budapešti (2002), St. Malo (2003), Ženevě (2005), Maastrichtu (2006), Göteborgu (2008) a Sarajevu (2009). Konference MIE 2011 se uskuteční v Oslu. Tradičně nebyly konference MIE organizovány v roce, kdy Mezinárodní asociace medicínské informatiky (IMIA)

organizuje své konference MEDINFO, které jsou pořádány jednou za tři roky [1]. EFMI také započala novou řadu setkání: Konference se speciálním tématem (STC). Koncept STC má následující prvky: a) organizací by měla být pověřena členská společnost, eventuálně v kombinaci se svou výroční schůzí; b) nedílnou částí STC by mělo být setkání Rady EFMI; c) téma konference by mělo být definováno organizující členskou společností; d) v souvislosti s tématem by měly být zapojeny relevantní pracovní skupiny EFMI; e) příspěvky by měly být hlavně zvané; STC jsou malé, dvoudenní konference se 100+ účastníky.

První konference STC se uskutečnila v roce 2001 v Bukurešti/Rumunsku, potom v roce 2002 v Nikosii/Kypru, v roce 2003 v Římě/Itálii, v roce 2004 v Mnichově/Německu, v roce 2005 v Aténách/Řecku, v roce 2006 v Temešváru/Rumunsku, v roce 2007 v Brijuni/Chorvatsku, v roce 2008 v Londýně/Velké Británii a v roce 2009 v Antalyi/Turecku. V roce 2010 se STC uskuteční v Rejkjavíku/Islandu a v roce 2011 v Ptuj/Slovensku. Sborník těchto konferencí byl většinou publikován nakladatelstvím Springer v sérii „*Lecture Notes in Medical Informatics*“ a nakladatelstvím IOS Press v sérii „*Studies in Health Technologies and Informatics*“. Výběr nejlepších příspěvků z konferencí MIE byl publikován ve speciálním vydání časopisu International Journal of Medical Informatics a bude publikován v časopise Methods of Information in Medicine, zatímco speciální sborník konference STC je publikován nakladatelstvím IOS Press jako sbírka recenzovaných článků.

2. Přehled časopisů zaměřených na medicínskou informatiku v Evropě

Vědecká a výzkumná produkce v oblasti medicínské informatiky v Evropě narůstá jak v kvantitě, tak i v kvalitě. Podporovat vysoce kvalitní výzkum je hlavním cílem EFMI [1], [4], [6]. EFMI má dva vysoce respektované oficiální časopisy, International Journal of Medical Informatics (bývalý název International Journal of Bio-Medical Computing), jehož současnými editory jsou Charles Safran a Jan Talmon a Methods of Information in Medicine se současným editorem Reinholdem Hauxem. EFMI také vydává několik specializovaných oficiálních časopisů, které pokrývají spektrum podoborů medicínské informatiky. Starobylou metodou jak šířit vysoce kvalitní vědecký výzkum je pomocí sborníků z každoroční konference MIE. EFMI také hraje důležitou roli ve vzdělávání a v harmonizaci podoborů medicínské informatiky organizováním konferencí STC. Jak sborníky z MIE, tak i z STC jsou recenzované a vydávané nakladatelstvím IOS Press (Amsterdam, Nizozemí). Sborníky z MIE jsou většinou zařazeny v Medline/PubMed. Většina časopisů medicínské informatiky je vydávána v národních jazycích, ale hodně z nich obsahuje anglické souhrny nebo obsahují kombinované články, které jsou publikovány v národním a anglickém jazyce. EFMI v podstatě poskytuje vysoce efektivní prostředky pro šíření výzkumu z oblasti medicínské informatiky, který je prováděn v Evropě, po celém světě. Kromě sborníků mají editoři časopisů možnost udělat víc pro diseminaci znalostí a zvýšení povědomí o našem oboru.

3. Analýza publikační aktivity v rámci členů EFMI

Většinu informací o EFMI lze najít na oficiálních webových stránkách (www.efmi.org). Navíc, většina národních společností má své vlastní webové stránky, které poskytují užitečné informace o publikacích a setkání v členských státech EFMI. Naše analýza je založena na faktech, generovaných z databáze EFMI, která je uložena na webových stránkách EFMI a která byla posbírána od národních představitelů členských států EFMI.

Když do Medline zadáme hledaný termín „EFMI“ a „European Federation for Biomedical Informatics“, objeví se pouze 21 článků 18 v časopisech vztahujících se k medicínské informatice: Informatics in Primary Care (6), International Journal of Medical Informatics (3), Methods of Information in Medicine (2), Studies in Health Technologies (6) a 3 v neinformatických/ne EFMI časopisech. Autory těchto článků jsou: Blobel (2), de Lusignan (6), France (1), Horsch (1), McKeon (1), Masic (2), Mihalas (2), Naszłady (1), Tallberg (1), Talmon (1), Trpišovský (1), van Goor (1) a Zvárová (1). (Pro srovnání) bylo provedeno hledání termínu „IMIA“ v Medline a bylo nalezeno 150 článků; při vyhledávání termínu „AMIA“ bylo nalezeno 244 článků. Většina EFMI článků se vztahuje k aktivitám pracovních skupin nejvíce vyčnívá PCI [5], [6] a relativně málo iniciativ je na úrovni federace. Detailní analýza webových stránek EFMI ukázala, že pouze 14 z 31 členů má vyjmenované své oficiální vědecké nebo profesní časopisy medicínské informatiky nebo bulletiny (Tab. 1). Časopisy medicínské informatiky mají své články publikovány v angličtině a tři jsou publikovány v národním jazyce. Existují tři časopisy, které jsou publikovány pouze jako elektronické.

Frekvence publikací s ISSN se mění: jeden časopis je vydáván jednou za dva roky, dva jsou čtvrtletní, jeden je vydáván jednou za dva měsíce, pět časopisů vychází 2-3krát za rok, dva časopisy jednou za měsíc a jeden časopis devětkrát za rok. Medicínská informatika je ústředním tématem všech těchto časopisů, ale některé z nich se zaměřují i na jiné podobory a příbuzné disciplíny. Většina z časopisů medicínské informatiky jsou mezinárodní publikace, i když jsou omezeny na národní uvádění do oběhu. Dva časopisy jsou financovány ministerstvem zdravotnictví a/nebo jinými veřejnými institucemi, zatímco ostatní jako způsob financování využívají předplatné.

4. Vyhlášení mise pracovní skupiny editorů EFMI časopisů

Vedoucí redaktori časopisů medicínské informatiky jsou vyzýváni, aby vytvořili společné vyhlášení mise a zamysleli se na cílem a účelem všech EFMI časopisů.

Věříme, že společné vyhlášení podpoří diseminaci vědeckých znalostí a zvýší povědomí o medicínské informatice jako vědním oboru. Navíc, máme v plánu publikovat a uveřejnit dokument, který bude vyjadřovat základní principy, na kterých by se měli editoři časopisů medicínské informatiky domluvit. Budou identifikovány společné cíle a bude odsouhlasen rámec a standard článků, které by měly být publikovány v EFMI časopisech medicínské informatiky. Důvody k ustanovení vyhlášení jsou následující:

- Povýšení vydavatelské dokonalosti prosazováním použití standardů a doporučených postupů (doporučení Editorů mezinárodní komise medicínských časopisů; Komise publikační etiky; další relevantní dokumenty [7]);
- Zvýšení kvality vědeckých studií publikovaných v časopisech medicínské informatiky zavedením stejných standardů pro hodnotící studie v medicínské informatice (doporučení již byla přijata jako oficiální dokument EFMI, ale zatím nebyly provedeny žádné další kroky k zajištění jejich používání);
- Povýšení vědecké vydavatelské dokonalosti prosazováním standardů ohledně formátu článku (příspěvek ve formátu „camera-ready“; Vancouverský referenční styl);
- Zlepšení šíření vědeckých znalostí skrz oblasti medicínské/zdravotnické informatiky (mělo by být propagováno pochopení a šíření EFMI výzkumu, výuky, doporučení klinické praxe vzemích EU);
- Zvýšení spolupráce mezi editory časopisů medicínské informatiky a naplánování pravidelných setkání pracovní skupiny;
- Prozkoumání potenciálu využití sdílených recenzentů, sdílení informací o recenzním procesu a sdílení odborných znalostí; technických redakčních informací, zkušeností, iniciativ, publikačních zdrojů a nástrojů (jako otevřený zdroj Open Journal System nebo komerční/nakladatelská řešení);
- Povzbuzení článků, které nejsou vhodné pro časopisy medicínské informatiky, aby byly publikovány někde jinde a podporovat komunikaci s editory těchto časopisů;

Tab. 1. - EFMI member countries medical informatics journals.

EFMI stát	Název časopisu	Vydavatel	Vedoucí redaktor	ISSN	Rok založení	Počet vydání za rok	Podání příspěvku	Webové stránky	Jazyk	Oblast	Recenze	Financování
Rakousko	Artificial Intelligence in	Elsevier	Klaus-Peter Adlassnig	1386-5056	1989	9	email	ano	angličtina	MI, další	mezinárodní	předplatné
Bosna a Hercegovina	Acta Informatica Medica	AVICENA, Sarajevo	Izet Masic	0353-8109	1993	4	email	ano	angličtina	MI, další	mezinárodní	předplatné
Chorvatsko	Bilten HDMI (bulletin CSMI)	CSMI	Kristina Fišter	1330-0253	1992	2	email	ano	chorvatština	MI, další	ne	členství ve společnosti
Česká republika	Medicinska informatika		Josipa Kern	1330-1799	1993	2			angličtina/chorvatština	MI	národní	Ministerstvo vědy, inzerce
Česká republika	Lékař a Technika	SBMILI	Karel Roubik	0301-5491	1970	4	email	ano	čeština	MI, biomedicínské	národní	veřejné instituce, předplatné
	European Journal for Biomedical	EuroMISE	Jana Zvarova	1801-5603	2005	2	-	ano	angličtina, další	Biomedicínská informatika	mezinárodní	jiné
Dánsko	Newsletter	Springer-Verlag	-			-		ano	dánština/angličtina	MI	národní	veřejné instituce, předplatné
Francie	Informatique et Sante	Urban &	Patrice Degoulet			2	email	ano	francouzština/angličtina	MI	národní	veřejné instituce, předplatné
Německo	Biometrie, Informatik und Epidemiologie in	Fisher Verlag				-	email	ano	němčina	MI, další	národní	veřejné instituce, předplatné
Irsko	Newsletter		-									
Izrael	ILAMI journal and The Health Screen	SMIA	-	-		2				MI, další		
Slovensko	Informatica Medica Slovenica	SEIS	Gaj Vidmar	1318-2129	1993	2	email	ano	angličtina, slovinština, další	MI, biostatistika, další	mezinárodní, národní	společnost MI (SMIA)
Španělsko	IHS	Elsevier	-			6				MI, další		
Švédsko	Computer Methods and Programs in	UAMI	Torny Groth	0169-2607	1985	12	papírově/email/online	ano	angličtina	MI	mezinárodní	předplatné
Ukrajina	Clinical Informatics and Telemedicine	Radcliffe	Oleg Yu Mayorov	1812-7231	2003	2-3	email	ano	ukrajinština, angličtina	MI, další	mezinárodní	veřejná instituce
Velká Británie	Informatics in Primary Care	Publishing	Simon de Lusignan	1476-0320	1993	4	papírově,online	ano	angličtina	PHC, klinická informatika	mezinárodní	PHCG BCS, předplatné
EFMI	International Journal of Medical Informatics	Elsevier	Charles Safran, Jan Talmon	1386-5056	1961	12	papírově/online	ano	angličtina	MI	mezinárodní	předplatné
	Methods of Information in	Schattauer	Reinhold Haux	0026-1270	1962	6	papírově/online	ano	angličtina	MI	mezinárodní	předplatné

- Poskytování společné řeči, když se objeví problematiky společného zájmu a zvýšení spolupráce mezi národními společnostmi a orgány EFMI;
- Povýšení evropských iniciativ v podněcování publikační činnosti a prvotřídního výzkumu.

Poděkování

Autoři by rádi poděkovali za podporu Janu Talmonovi z Maastricht University, Care and Public Health Research Institute, Nizozemí.

Literatura

- [1] Masic I, Novo A.: History of medical informatics in Bosnia and Herzegovina. Acta Informatica Medica.2007;15(1):49-61.
- [2] Collen M. F.: Fifty years in medical informatics. IMIA Yearbook of Medical Informatics. Methods of Information in Medicine. 2006;45(Suppl. 1):S174 -9.
- [3] Zvárová J.: Symposium on "Biomedical Informatics and Biomedical Statistics Education". Methods of Information in Medicine. 2006;45(3):281-2.
- [4] Zvárová J.: Biomedical Informatics Research and Education in the EuroMISE Center. IMIA Yearbook of Medical Informatics. Methods of Information in Medicine. 2006;45(Suppl. 1):S166-73.
- [5] de Lusignan S.: Introduction to the Primary Care Informatics Working Group of the European Federation for Medical Informatics (EFMI). Informatics in Primary Care. 2003;11(3):175-6.
- [6] de Lusignan S., Teasdale S.: Achieving benefit for patients in primary care informatics: The report of an international consensus workshop at MedInfo 2007. Informatics in Primary Care.2007;15(4):255-61.
- [7] Miller R. A., Groth T/, Hasman A/, Safran C., Shortliffe E. H., Haux R., McGray A. T.: On exemplary scientific conduct regarding submission of manuscripts to biomedical informatics journals. Methods of Information in Medicine.2006;45(1):1-3.

Kontakt

Izet Masic
 Medical Faculty
 University of Sarajevo
 Cekalusa 90/4
 71000 Sarajevo
 Bosnia and Herzegovina
 e-mail: imasic@lol.ba

Robustní analýza obrazu obličeje pro genetické aplikace

Jan Kalina¹

¹Centre of Biomedical Informatics, Institute of Computer Science AS CR, Prague, Czech Republic

Souhrn

Tento článek se zabývá automatickou lokalizací objektů (očí, úst) ve dvourozměrných (2D) černobílých obrazech obličejů. Je motivován praktickým problémem v genetice člověka a výstup lokalizace objektů v dané databázi obrazů je zapotřebí pro řešení dalších úloh v genetickém výzkumu. V článku se aplikuje robustní filtr na obrazy s cílem odstranit šum. Hlavní metodou jsou šablony. Ústa a obě oči se lokalizují současně za použití váženého Pearsonova korelačního koeficientu nebo jeho robustní analogie založené na robustních regresních metodách. V databázi s 212 obrazy obličejů tato metoda správně nalezne ústa a oči ve 100 % případů. Také robustní korelační koeficient založený na regresní metodě nejmenších vážených čtverců lokalizuje ústa a oči ve 100 % obrazů uvažované databáze. Článek studuje robustní aspekty této metody vzhledem k otočení, šumu, okluzi a asymetrii v obraze. Současná lokalizace úst i obou očí je invariantní vůči libovolnému otočení obličeje. Tato studie využívá speciální vlastnosti daných obrazů obličejů vzhledem k očekávanému použití v genetických aplikacích.

Klíčová slova: lokalizace objektů, šablony, detekce očí a úst, robustní korelační analýza, redukce šumu

1. Úvod

Hlavním cílem této práce je automatická lokalizace orientačních bodů v obrazech obličejů motivovaná výzkumem v genetice člověka. Pracujeme s databází obrazů, které pocházejí z Ústavu genetiky člověka na Univerzitě Duisburg-Essen v SRN (projekty BO 1955/2-1 a WU 314/2-1 Německé akademie věd DFG). Tato databáze obsahuje 212 černobílých obrazů velikosti 192*256 pixelů, kde každý obraz obsahuje právě jeden obličej. Každý obraz zobrazuje jinou osobu. Obličeje mají zhruba stejnou velikost, zatímco vykazují drobné otočení ve frontální rovině. Naším

cílem je nalézt robustní řešení vůči rotaci, okluzi, šumu v obraze a asymetrii obličeje, které bude mít zároveň jasnou interpretaci.

Šablony jsou vhodnou metodou pro detekci objektů v černobílých obrazech. Šablona vystihuje modelový, typický nebo ideální tvar uvažovaného objektu. Položí se na libovolnou možnou pozici v obraze, přičemž se měří podobnost mezi šablonou a odpovídající částí obrazu tak, že se porovná intenzita šedi v každém pixelu šablony s intenzitou šedi odpovídajícího pixelu v obraze. [1] udává výčet referencí s aplikacemi šablon do detekce obličejů a rozpoznávání obličejů. Běžné metody analýzy obrazu se ovšem konstruují jako kaskády takových jednoduchých klasifikátorů [2]. Naproti tomu jasná interpretace a studium robustních vlastností [3] je možné spíše tehdy, pokud ignorujeme obvyklé postupy pro redukci dimenze a extrakci příznaků.

[4] aplikuje šablony na datový soubor obrazů očních duhovek 64 osob s úlohou přiřadit novou duhovku správné osobě z trénovací databáze. Tato metoda maximalizuje vzájemnou (Shannonovu) informaci jakožto míru podobnosti mezi šablonou a červeným kanálem barevného obrazu. [5] uvažuje šablony jen jako prvky hierarchické struktury stromového typu. Složitější algoritmy analýzy obrazu [2] kombinují metody matematiky, statistiky a informatiky s důrazem na vysokou výpočetní rychlost spíše než výhodné teoretické vlastnosti [3]. V poslední době se stalo v analýze obrazu zvykem nahrazovat pixely buď záplatami (patches) nebo příznaky (features). Záplatami [6] se rozumí homogenní oblasti pixelů, přičemž přístup založený na záplatách analyzuje celý obraz jako množinu jednotlivých záplat. Naproti tomu příznaky typicky odpovídají hranám či objektům s heterogenitou nebo nespojitostí. [7] byl prvním článkem, který využil redukci dimenze nahrazením intenzit jednotlivých pixelů sadou příznaků. [8] a [9] jsou novější

práce o detekci člověka nebo obličeje za pomoci extrakce příznaků, která je zároveň robustní ke změnám osvětlení, pózy nebo výrazu v obličeji. [10] studuje varianční matice příznaků, protože variabilita je obecně robustní ke změnám osvětlení.

Pracujeme s databází černobílých obrazů, v nichž každému pixelu odpovídá intenzita šedi z intervalu [0,1]. Malé hodnoty jsou černé a velké hodnoty bílé. Obrazy jsou vyfotografovány za standardizovaných podmínek tak, že osoba sedí rovně čelem proti fotoaparátu a dívá se přímo do něj. Ústav genetiky člověka usiloval o maximální standardizaci daných obrazů a z toho důvodů opakovaně fotografoval osoby tak, aby výsledná databáze neobsahovala obličeje se zavřenými očima, vlasy překrývající obličej či další rušivé efekty. Přesto jsou některé obličeje otočené ve frontální rovině o malý úhel tak, že oči neleží přesně vodorovně. Nicméně databáze neobsahuje obrazy s trojrozměrnou rotací.

Ústav genetiky člověka pracuje na zajímavých problémech v genetickém výzkumu s využitím obrazů obličejů [11]. Ambicí tohoto výzkumu je automatická klasifikace genetických syndromů z obrazu obličeje; dále studium souvislostí mezi genetickým kódem a velikostí i tvarem obličejových částí; a konečně vizualizace obličeje jen z hodnot biometrických veličin [12]. Obrazy pacientů mohou být klasifikovány do jedné z 10 skupin podle genetické deformace, která se projevuje v dysmorfii obličeje. Pro různé syndromy se úspěšnost této automatické klasifikace pohybuje mezi 75 % a 80 %, což se považuje za poměrně spolehlivý výsledek. Lokalizace orientačních bodů v obličeji je v každém případě prvním krokem všech těchto studií, i když není sama o sobě jejich primárním cílem.

Orientačními body jsou prominentní části obličeje, například koutky očí či úst, prostředek horního nebo spodního okraje rtů nebo význačné body v nosních dírkách a obočí. Genetici používají dvě metody pro nalezení 40 orientačních bodů v každém obličeji. Jednou možností je manuální identifikace, kterou pečlivě a přesně provádí zkušený antropolog. Jako druhou možnost ústav používá automatickou metodu založenou na algoritmu [13], kterou nyní popíšeme.

Algoritmus začíná manuální lokalizací sady 40 orientačních bodů v trénovací databázi 83 obrazů obličejů. Sada těchto orientačních bodů je jako celek položena na všechny možné pozice v obraze jako jedna velká šablona s pevnými vzdálenostmi mezi jednotlivými orientačními body. Pak se provede dvourozměrná Gaborova waveletová transformace s různými hodnotami dvourozměrného parametru pro měřítko na všechny obrazy trénovací databáze a také na nový obraz, v němž se mají lokalizovat orientační body. Porovnají se waveletové koeficienty v každém orientačním bodě trénovacího obrazu s waveletovými koeficienty v odpovídajících pixelech nového obrazu. Přitom můžeme chápat tyto sady Gaborových waveletových koeficientů jako mnohorozměrné šablony. Nakonec se spočítá korelační koeficient mezi vektory waveletových koeficientů a tyto se sečtou přes všech 40 orientačních bodů. Tato suma je pak mírou podobnosti mezi trénovacím a novým obrazem. Nicméně se ukazuje, že tento postup je příliš citlivý vůči malým otočením obličeje.

Cílem naší práce je hledat ústa a oči v obrazech obličejů za pomoci šablon. Navrhujeme zde algoritmus pro lokalizaci každého oka zvlášť a také současně obou očí i úst za pomoci šablon. Studujeme robustní vlastnosti popsaných metod, například vůči šumu či otočení obrazu. Článek popisuje počáteční redukci šumu v obraze za pomoci robustního filtru. Hlavní metodou jsou šablony použité na lokalizaci očí a na současnou lokalizaci obou očí i úst. Důraz klademe na robustní aspekty těchto metod a robustní analogie korelačního koeficientu.

2. Metody

Nejprve provedeme redukci či odstranění šumu z obrazů. Pak popíšeme metodu pro lokalizaci očí v obrazech a následně současnou lokalizaci obou očí a úst. Nakonec porovnáme různé metody robustní korelační analýzy pro tutéž úlohu současně lokalizace obou očí a úst. Je třeba připustit, že při práci s originálními obrazy se ztrácí výhoda postupů založených na příznacích (kapitola 1), které umožňují pracovat s různými velikostmi obrazů. Náš přístup je však invariantní vůči rotaci a navíc má i jasnou interpretaci. Výsledky jsou shrnuty v kapitole 3.

2.1 Vyhazení šumu

Redukce či vyhlazení šumu, filtr nebo robustifikace jsou synonyma pro transformaci často používanou pro potlačení vlivu šumu v obrazech. Shrneme zde argumenty ve prospěch používání takové transformace a popíšeme postup založený na regresních odhadech metodou nejmenších usekaných čtverců nebo nejmenších vážených čtverců. Důležité je, aby i při potlačení šumu v obraze byly stále dobře rozeznatelné jednotlivé obličejové části. Přitom při vyhlazení šumu je třeba odstranit jeden artefakt v naší databázi obrazů, kterým je odraz světla na různých místech v očích. Tento rušivý efekt byl způsoben při fotografování na Ústavu genetiky člověka. Popíšeme nyní vysoce robustní odhady v lineárním regresním modelu, které zároveň použijeme v kapitole 2.4 pro definici robustního korelačního koeficientu.

[14] popisuje filtry (dvourozměrné operátory) pro vyhlazení šumu a dává přednost useknutému průměru a jiným L-odhadům [15] před mediánem. [16] definuje M-korelační koeficient založený na teorii M-odhadů, který kombinuje robustnost a eficienci pro Gaussovský bílý šum, a aplikuje ho do analýzy obrazu za pomoci šablon. Vyhazení šumu má své aplikace i v obrazech v molekulární genetice [17] či alternativně [18] za pomoci robustních statistických metod.

Definujme nyní robustní odhady, které použijeme na vyhlazení šumu v obrazech naší databáze. Uvažujme lineární regresní model ve tvaru

$$Y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} + e_i, i = 1, 2, \dots, n, \quad (1)$$

který můžeme přepsat v maticovém značení jako $Y = X\beta + e$. Regresní odhad metodou nejmenších vážených čtverců (least weighted squares, LWS) [19] je jednou z robustních regresních metod s vysokým bodem selhání. Vyžaduje specifikaci velikostí nezáporných vah w_1, w_2, \dots, w_n . Tyto váhy jsou ovšem přiřazeny jednotlivým pozorováním až po provedení permutace, která se určí teprve implicitně v průběhu výpočtu tohoto odhadu. Tato permutace závisí na hodnotách reziduí $u_1(b), \dots, u_n(b)$, které přísluší konkrétní hodnotě $b = (b_0, b_1, \dots, b_p)^T$ odhadu vektoru regresních parametrů β , kde

$$u_i(b) = Y_i - b_0 - b_1 x_{i1} - \dots - b_p x_{ip}, i = 1, 2, \dots, n. \quad (2)$$

Typicky se volí váhy w_1, w_2, \dots, w_n jako nerostoucí posloupnost. Pokud označíme uspořádané hodnoty druhých mocnin reziduí pomocí

$$u_{(1)}^2(b) \leq u_{(2)}^2(b) \leq \dots \leq u_{(n)}^2(b), \quad (3)$$

pak je odhad β metodou LWS definován jako

$$\hat{b}_{LWS} = \underset{b}{\operatorname{argmin}} \sum_{i=1}^n w_i u_{(i)}^2(b)$$

Oblíbenou volbou vah jsou lineárně klesající váhy [20] $w_i = 1 - (i-1)/n$, $i = 1, \dots, n$; jinou možnost představuje dvoukroková procedura [21] pro výpočet adaptivních vah, která umožňuje automaticky určit také velikosti vah.

Nejmenší vážené čtverce kombinují robustnost s eficiencí [21]. Rychlý přibližný algoritmus pro výpočet LWS se získá jako vážená analogie algoritmu [22]. [23] studuje metodu nejmenších usekaných čtverců (least trimmed squares, LTS), která je speciálním případem LWS regrese s vahami rovnými pouze 1 nebo 0. Přitom se musí určit počet pozorování h , které dostanou váhu rovnou 1.

Použijeme tyto robustní odhady v kontextu vyhlazení obrazu. Jejich výhodou je vysoký bod selhání [15], který zajistí velkou odolnost vůči odlehlým pozorováním. Na místo lineárního regresního modelu zde stačí uvažovat jen model s parametrem polohy. Odhad metodou nejmenšího mediánu čtverců (least median of squares, LMS) [23] je ekvivalentní průměru nejkratší poloviny dat. Obdobně LTS odhad odpovídá průměru takové poloviny dat (nebo skupiny nějakých $h < n$ pozorování), která má nejmenší rozptyl. LWS odhad pak je váženým průměrem dat s takovou permutací vah, která vede k nejmenšímu váženému rozptylu dat [20].

Uvažujeme kruhové okolí každého pixelu a spočítáme odhad metodou LMS, LTS nebo LWS z intenzit šedi jednotlivých pixelů. Přitom se ztrácí informace o souřadnicích každého pixelu. Medián vede jen ke slabým výsledkům, protože odstraňuje kontrast a poskytuje poněkud našedlý výsledný obraz. Výkon metod LTS a LWS je mnohem přesvědčivější, obzvláště když zvolíme malý poloměr pro kruhové okolí každého pixelu. Výsledek není příliš ovlivněn volbou h pro LTS metodu nebo volbou vah pro LWS metodu.

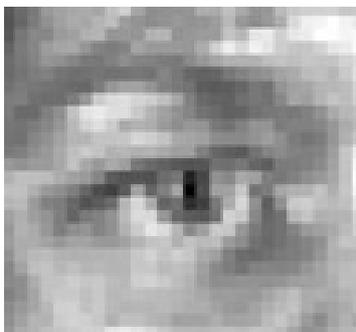
Zvolili jsme LTS filtr a použijeme jej na každý obraz v naší databázi. Uvažujeme 4 přímé sousedy každého pixelu a spočítáme LTS odhad s hodnotou $h=3$, která se spočítá jako průměr nejkratší trojice získané z uvažovaných 5 hodnot uspořádaných od nejmenší po největší. Takto se odstraní nejen možné odlehlé hodnoty, ale také světlo odrážené v očích. Dále jsme zkoumali rezidua této transformace s cílem zjistit její efekt. Hodnoty reziduí velké v absolutní hodnotě ukazují na velký lokální efekt tohoto vyhlazení. Velký efekt je nápadně přítomen na hraničních oblastech v obraze jako na okraji vlasů či mezi vlasy a oblečením, mezi vlasy a pozadím nebo na hranici nosních dírek.

2.2 Lokalizace očí

Za pomoci šablon budeme hledat v každém obraze zvláště levé a pravé oko. Oblíbenou volbou šablony je průměr [24], [25], a proto zde použijeme průměr očí jako šablonu pro lokalizaci očí. Zkonstruovali

jsme sadu 6 očních šablon pro pravé oko a jejich překlopení v osové symetrii pro levé oko. Šablony jsme získali jako průměr skutečných stejně velkých očí různých osob; jednu z nich ukazuje Obrázek 1. Šablony mají různé velikosti mezi 26×28 pixelů a 36×30 pixelů. Pearsonův korelační koeficient r nebo vážený Pearsonův korelační koeficient r_w s vhodnými vahami použijeme jako míru podobnosti mezi šablonou a obrazem.

Všechny šablony obsahují obočí, které by případně mohlo zkomplikovat lokalizaci očí. Problém však vyřešíme tak, že oblasti obočí přiřadíme menší váhy pro vážený korelační koeficient. Obrázek 1 ukazuje šablonu pro oko, která dává nejlepší výsledky při lokalizaci pravého oka;



Obr. 1. Šablona pro pravé oko.



Obr. 2. Korelační koeficient mezi daným obrazem a šablonou pro pravé oko. Oblasti v obraze s velkou hodnotou korelačního koeficientu jsou zobrazeny černě.

v tomto případě jsme použili jedinou šablonu spolu s váženým korelačním koeficientem s radiálními vahami.

Spočítejme nyní korelační koeficient mezi jedním obrazem z databáze a oční šablonou z Obrázku 1. Hodnoty korelačního koeficientu ukazuje Obrázek 2, v němž černé oblasti mají velkou hodnotu korelačního koeficientu. Tyto oblasti obsahují obě oči a také ústa či část vlasů. Zajímavé je, že pravé oko dané osoby má větší korelační koeficient s jednou šablonou pro levé oko než s kteroukoli šablonou pro pravé oko. V algoritmu využijeme tento jev následujícím způsobem. Položíme všech 12 šablon na libovolnou možnou pozici v obraze a nerozlišujeme mezi šablonami pro levé a pravé oko. Nejprve najdeme oblast s největším korelačním koeficientem přes všechny šablony. To je jedna kandidátní oblast, která může obsahovat oko; řekněme jí kandidátní oko. Dále uvažujeme celý obraz s výjimkou tohoto kandidátního oka a jeho nejbližšího okolí. V takovém obraze najdeme znovu oblast s největším korelačním koeficientem s libovolnou oční šablonou. Ta bude druhým kandidátním okem.

2.3 Lokalizace očí

Zde si klademe za cíl lokalizovat současně ústa a obě oči. Najdeme tedy nejprve kandidátní ústa za pomoci šablon a hledáme oči v určité oblasti nad kandidátními ústy. Používáme 6 šablon pro pravé oko, jejich překlopení v osové symetrii a 7 symetrických šablon pro ústa. Každou ze šablon pro ústa necháme najít několik oblastí, které mají vážený korelační koeficient s libovolnou šablonou pro ústa větší než určitou hodnotu. Ta může být různá v různých obrazech. Vždy však hledáme alespoň tři různé oblasti, které nejsou svými přímými sousedy. Tyto oblasti považujeme za kandidátní ústa. Nejprve najdeme oblasti s váženým korelačním koeficientem r_w přesahujícím hodnotu 80 % maximální hodnoty r_w , které je dosaženo v daném obraze. Tato mez se může případně snížit, aby se podařilo najít alespoň troje kandidátní ústa.

Různé šablony pro ústa vedou k různým kandidátním ústům. Pak hledáme oči v takových oblastech, které odpovídají jednotlivým kandidátním ústům. Tyto oblasti se nacházejí od 20 do 55 pixelů nad kandidátními ústy a považujeme je za kandidátní oblasti pro oči. Umístíme každou ze 6 šablon pro oči a její překlopení v osové symetrii na libovolnou možnou pozici v této kandidátní oblasti pro oči a spočítáme vážený korelační koeficient mezi šablonou a odpovídající stejně velkou oblastí. Zároveň rozdělíme kandidátní oblast pro oči na dvě části podél myšlené svislé čáry (Obrázek 2), která rozděljuje ústa napůl. Všechny šablony pro oči a také jejich překlopení v osové symetrii jsou dále použity v obou částech kandidátní oblasti pro oči. Hledáme oblast s největším váženým korelačním koeficientem s libovolnou šablonou zvlášť v levé a pravé části. Tento postup se opakuje pro několik kandidátních úst.

Navíc používáme podmínku na vzájemnou vzdálenost obou očí s mezemi mezi 25 a 42 pixely. Tyto meze jsme získali empiricky tak, aby je skutečné oči splňovaly dokonce v obrazech s mírně odlišnou velikostí nebo menší rotací obličeje.

Pro konkrétní kandidátní ústa najdeme kandidátní oblast pro pravé a levé oko. Jestliže nesplňují podmínku na vzdálenost

mezi očima, pokračujeme v hledání takové oblasti, která má největší vážený korelační koeficient s libovolnou oční šablonou mezi všemi zbývajícimi oblastmi v kandidátní oblasti pro oči. Taková oblast se středem v kandidátní oblasti pro oči je jedním kandidátním okem, zatímco druhé kandidátní oko má největší vážený korelační koeficient s libovolnou šablonou a musí mít zároveň svůj střed v opačné části kandidátní oblasti. Znovu zkontrolujeme podmínku na vzdálenost očí a je-li třeba, opakujeme tyto kroky, až najdeme dvě kandidátní oči, které splňují danou podmínku na vzdálenost. Tato metody výrazně zlepšuje výsledky pro otočené obličeje (kapitola 3.4).

Základní myšlenkou je sečíst tři hodnoty váženého korelačního koeficientu odpovídající ústům a oběma očím. Začneme s jednou z těch oblastí, které mají velkou hodnotu váženého korelačního koeficientu s libovolnou šablonou pro ústa. Označme největší hodnotu mezi těmito hodnotami váženého korelačního koeficientu mezi šablonou a oblastí obrazu se středem v daném pixelu jako r_{w1} . Umístíme všechny oční šablony na všechny možné pozice do kandidátní oblasti pro oči tak, aby střed byl vždy v její levé části. Hledáme oko jen v této levé části a označíme největší hodnotu váženého korelačního koeficientu přes všechny oční šablony jako r_{w2} . Podobným způsobem umístíme všechny oční šablony do pravé části kandidátní oblasti a označíme největší vážený korelační koeficient pomocí r_{w3} .

Uvažujme koeficient

$$r_w^* = (r_{w1} + r_{w2} + r_{w3}) \cdot I(r_{w1} > 0) \cdot I(r_{w2} > 0) \cdot I(r_{w3} > 0), \quad (4)$$

kde 1 označuje indikátor. (4) ignoruje záporné hodnoty váženého korelačního koeficientu. Největší hodnotu r_w^* spočteme pro různá kandidátní ústa.

2.4 Robustní korelace

Použijeme robustní verzi korelačního koeficientu pro společnou lokalizaci obou očí a úst v obrazech. Uvažujeme koeficienty založené na robustní odhadech v lineárním regresním modelu. Odstranění některých pixelů odpovídá myšlence, že některé pixely jsou redundantní a ignorování velkých oblastí pixelů může proto připomínat postupy založené na záplatách [26].

Výpočet provedeme tak, že nejprve převedeme šablonu a příslušnou stejně velkou část obrazu na vektory, spočítáme LWS regresi, v níž obraz vysvětlujeme jako odezvu šablony, pak uvažujeme váhy určené LWS metodou a spočítáme vážený korelační koeficient mezi obrazem a šablonou s těmito vahami. Uvažujeme lineárně klesající váhy nebo adaptivní váhy [21]. Korelační koeficient založený na LTS se definuje obdobně. Zde jsme zvolili, že požadujeme odříznout 20 % pixelů ($h = 0.8n$), které budeme považovat za odlehle hodnoty.

Tab. 1. Procenta správných výsledků při lokalizaci očí za pomoci 6 šablon a při společné lokalizaci úst a očí za pomoci 7 šablon pro ústa a 6 šablon pro oči. Vážený Pearsonův korelační koeficient je použit s radiálními vahami. Originální obrazy jsme modifikovali změnou velikosti o 10 % nebo otočením obrazu o $+10^\circ$ nebo -10° .

Šablony	Originální obraz	Zmenšený obraz	Otočený obraz
Oči, r	1.00	0.80	0.50
Oči, r_w	1.00	0.97	0.86
Oči a ústa, r	1.00	0.99	1.00
Oči a ústa, r_w	1.00	1.00	1.00
Oči a ústa (relaxovaná metoda), r	1.00	0.99	0.92
Oči a ústa (relaxovaná metoda), r_w	1.00	0.99	0.87

3. Výsledky

3.1 Lokalizace očí

Sada 6 šablon pro oči (kapitola 2.2) lokalizuje obě oči správně ve 100 % obrazů uvažované databáze za použití Pearsonova korelačního koeficientu jakožto míry podobnosti mezi obrazem a šablonou. To znamená, že výsledek lokalizace odpovídá skutečným očím ve všech daných obrazech. Správné výsledky ve 100 % případů se získají také s váženým korelačním koeficientem s radiálními vahami, které se definují jako nepřímo úměrně vzdálenosti každého pixelu od středu šablony. Tyto váhy zároveň zajistí lepší výsledky z hlediska robustních vlastností (kapitola 4).

Tabulka 1 shrnuje výsledky lokalizace obou očí zvlášť za pomoci 6 šablon a jejich překlopení v osové symetrii. Dále dává výsledky pro další situace popsané později, u nichž porovnáva výstup získaný s korelačním koeficientem r a jeho váženou obdobou r_w s radiálními vahami.

3.2 Lokalizace očí a úst

Metoda (4) pro společnou lokalizaci obou očí a úst (kapitola 2.3) správně lokalizuje ústa a obě oči ve 100 % obrazů naší databáze. Tabulka 1 prezentuje výsledky této metody získané pomocí r a r_w s radiálními vahami a také výsledky získané relaxovanou verzí téže metody, která nepoužívá pravidlo omezující vzdálenost obou očí. Z výsledků plyne důležitost takového pravidla. Dále Tabulka 1 obsahuje výsledky získané pro obrazy s modifikovanou velikostí nebo rotací, které budou popsány v kapitole 3.4.

3.3 Robustní korelace

Pracujeme s databází s 212 obrazy a 7 šablonami pro ústa a 6 šablonami pro oči spolu s jejich překlopeními v osové symetrii. Několik robustních korelačních metod z kapitoly 2.4 dává výsledky současné lokalizace obou očí a úst správně ve 100 % případů. Tyto výsledky jsme získali pomocí korelačního koeficientu založeného na LTS s volbou $h=0.8n$, dále s korelačním koeficientem založeným na LWS s lineárně klesajícími vahami a také s adaptivními vahami. Typicky se odlehle hodnoty odříznuté metodou LTS nacházejí v okolí očí, při okraji obdélníkové šablony. Také korelační

koeficient založený na LWS přiřazuje malé hodnoty vah těmto pixelům při okraji šablony.

Pro srovnání uvažujeme také robustní korelační koeficient založený na useknutí tzv. hlavních proměnných $u = x + y$ and $v = x - y$, které navrhuje [27]. Tyto selhávají v naší aplikaci při lokalizaci úst. Důvod vysvětlíme na korelačním koeficientu mezi ústy a šablonou pro ústa. Odlehle hodnoty v proměnné u se totiž typicky nacházejí ve rtech, zatímco odlehle hodnoty v proměnné v odpovídají tvářím. Naproti tomu při výpočtu korelačního koeficientu mezi neústy a šablonou pro ústa se odlehle pozorování typicky nacházejí ve tvářích. Není jednoznačný vztah mezi odlehlymi hodnotami v hlavních proměnných a v původních datech, což neumožňuje správnou klasifikaci úst.

Shrňme, že jsme aplikovali různé robustní verze korelačního koeficientu včetně těch, které jsou založeny na robustní regresi s vysokým bodem selhání. Tyto dávají výsledky správně ve 100 % případů při současné lokalizaci obou očí a úst, stejně tak jako vážený korelační koeficient s radiálními vahami.

Tab. 2. Procenta správných výsledků získané s váženým Pearsonovým korelačním koeficientem s radiálními vahami při lokalizaci obou očí a úst současně. Metoda se aplikuje na originální obrazy stejně jako na obrazy otočené o libovolný úhel a obrazy drobně modifikované šumem, okluzí či asymetrií.

Obrazy	r_w , radiální váhy
Originální	1.00
Otočení o libovolný úhel	1.00
Šum	1.00
Okluze	1.00
Asymetrie obličeje	1.00

3.4 Robustnost popsané metody

V této kapitole popíšeme validační kroky, které ověří chování naší metody na současnou lokalizaci úst a obou očí. Ověříme metodu (4) na validační databázi a prozkoumáme její vlastnosti na obrazech otočených o libovolný úhel a konečně prostudujeme robustnost na numerické

studii tak, že modifikujeme originální obrazy dodatečně přidaným šumem, malou okluzí nebo asymetrií. Tabulka 2 shrnuje modifikace originálních obrazů, které jsme provedli.

Tato metoda se neučí žádné parametry přes databázi obrazů, na kterou se používá. Specifické vlastnosti naší konkrétní databáze ovlivní jen výběr šablony a také očekávané vzdálenosti mezi oběma očima a dále mezi ústy a očima. Přesto ověříme chování naší metody na validační databázi obrazů. Vyfotografovali jsme 30 náhodně vybraných studentů na Univerzitě Duisburg-Essen kompaktním digitálním fotoaparát. Standardizovali jsme podmínky, abychom získali fotografie se stejným osvětlením a vzdáleností obličeje od fotoaparátu. Některé vyfotografované obličeje jsou mírně otočené ve frontální rovině. Obličeje mají vždy neutrální výraz ve tváři. Transformovali jsme barevné fotografie z původní velikosti 2048*1536 pixelů na černobílé obrazy velikosti 266*200 pixelů. Pak velikost obličeje odpovídá velikosti obličejů v původní databázi. Použili jsme metodu z kapitoly 2.3 na lokalizaci úst a obou očí v obrazech. Tato metoda dává správné výsledky ve 100 % případů za použití 7 šablon pro ústa a 6 šablon pro oči. Nicméně považujeme tuto metodu za příliš citlivou vůči velikosti obrazu, a proto ji neověřujeme na další databázích obrazů.

Teoretické vlastnosti šablon a jejich robustnost k rotaci, okluzi či asymetrii jsou studovány v článku [3], z nějž vyplývá, že koeficient (4) má robustní vlastnosti ve stejných situacích. Naproti tomu výběrová influenční funkce [28] není omezená a tedy hodnota (4) je příliš citlivá vůči velkým (influenčním) vahám.

Vyšetříme nyní naši metodu pro otočené obrazy, na něž aplikujeme neotočené šablony. Lokalizace každého oka zvlášť nebo samotných úst nedává úspěšné výsledky v obrazech otočených o $\pm 10^\circ$. Dále se věnujeme současné lokalizaci obou očí a úst v obrazech otočených o $\pm 10^\circ$.

Za použití 6 šablon pro oči, jejich překlopení v osové symetrii a 7 šablon pro ústa spolu s radiálními vahami pro všechny šablony pro ústa i pro oči nalezneme ústa a obě oči správně ve 100 % obrazů naší databáze. Nicméně otočení obrazů o $\pm 20^\circ$ už způsobí, že naše metoda selhává asi v 50 % obrazů.

Pokud otočíme obličej o libovolný úhel s velikostí 10, 20, ..., 350 stupňů, je součet tří vážených korelačních koeficientů menší než v neotočeném obličejí s očima ve vodorovné poloze. Dobrou strategií je otočit obličej o několik různých úhlů a najít největší hodnotu koeficientu (4) pro každé otočení. Tímto způsobem lokalizujeme ústa a obě oči správně nezávisle na počátečním otočení daného obrazu. Konkrétně jsme otočili každý obraz o úhly s velikostí 0, 10, 20, ..., 180, ..., 350 stupňů. Největší hodnota r_w^* v každém z 212 obrazů se získá právě v takové poloze, kdy jsou oči ve vodorovné poloze. To vede ke správné lokalizaci úst a obou očí ve 100 % obrazů. Tyto 100 % správné výsledky pro obrazy otočené o libovolný úhel jsme získali také s korelačním koeficientem založeným na LWS s lineárně klesajícími vahami a s adaptivními vahami.

Metoda pro lokalizaci úst a obou očí je vybavena dvojitou ochranou proti případnému otočení obličeje. Za prvé jsou šablony robustní vůči otočení o úhel až do velikosti $\pm 10^\circ$. A za druhé koeficient r_w^* nabývá své největší hodnoty právě pro neotočený obličej. Proto se správné otočení obličeje nalezne v každém obličejí s libovolným otočením až do $\pm 180^\circ$ stupňů. Tabulka 1 shrnuje výsledky získané pomocí r a r_w s radiálními vahami v obrazech otočených o $\pm 10^\circ$ nebo zmenšených o 10 %.

Konečně vyšetřujeme robustní vlastnosti vzhledem k dalším nestandardním situacím. Šum s normálním rozdělením s nulovou střední hodnotou přidáný do obrazu nezávisle jednotlivým pixelům nepoškodí správné výsledky, pokud jeho rozptyl nepřekročí hodnotu $\sigma = (0.11)^2$. Ústa a obě oči se správně lokalizují také v obrazech otočených o libovolný úhel (Obrázek 3) a v obrazech, do nichž jsme dodatečně přidali šum s nulovou střední



Obr. 3. Lokalizace úst a obou očí. Pro konkrétní kandidátní ústa se hledají obě oči ve vyznačené relevantní kandidátní oblasti.

hodnotou a rozptylem σ^2 nepřekračujícím 0.01.

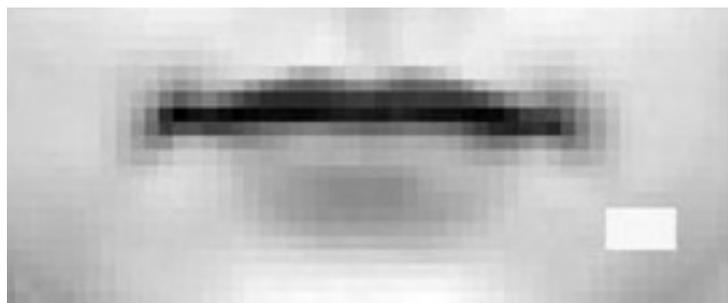
Okluzí rozumíme zakrytí části obrazu jiným obrazem nebo jinou abnormalitou. Provedli jsme studii, při níž jsme ústa v každém obrazu zakryli malou náplastí (Obrázek 5), abychom vyšetřili lokální citlivost naší metody. Intenzity šedi v obdélníku velikosti 3*5 pixelů jsme nastavili na hodnotu 1. Každá ústa v databázi modifikujeme tímto způsobem tak, že umístíme náplast vždy na stejnou pozici do pravého dolního rohu úst, kde je umístěn se vzdáleností od 7 do 9 pixelů pod středem úst a se vzdáleností od 16 do

20 pixelů napravo od středu úst. Příkladem jsou právě takto modifikovaná ústa v Obrázku 5. Ve 100 % obrazů s takto modifikovanými ústy jsme ústa a obě oči lokalizovali správně.

Další studie se věnuje vlivu asymetrie úst. Intenzity šedi v pravé polovině každých úst jsme navýšili o konstantu, kterou označme ϵ . Společná lokalizace úst a obou očí dává 100 % správné výsledky s hodnotou ϵ až do 0.15, což již představuje docela závažné zkreslení originálních úst. Zde selhává korelační koeficient i vážený korelační koeficient s radiálními vahami při lokalizaci samotných úst za pomoci 7 symetrických šablon pro ústa.



Obr. 4. Obrázek otočeného obličeje dodatečně doplněný šumem.



Obr. 5. Ústa modifikovaná okluzí malého rozměru (bílé překrytí originálního obrazu).

Ověřili jsme tedy, že současná lokalizace úst a očí se 7 šablonami pro ústa a 6 šablonami pro oči má požadované robustní vlastnosti vzhledem k malé rotaci, okluzi a asymetrii obrazu.

4. Diskuse a závěry

Navrhli jsme metodu pro současnou lokalizaci úst a obou očí v obrazech obličejů. Naše postupy jsou ušité na míru pro danou databázi obličejů s očekávaným použitím v genetických aplikacích. Šablony jsou koncepčně jednoduché a snadno interpretovatelné. Naše metoda je robustní k šumu v obrazech, drobné okluzi či asymetrii obličejů. Správné výsledky jsme získali pro libovolné počáteční otočení obličeje. Nevýhodou našeho postupu pro všeobecné použití ovšem je citlivost k velikosti obrazu.

Náš pokus aplikovat různé verze robustního korelačního koeficientu na problém lokalizace objektů v obraze vedl ke slibným výsledkům a můžeme proto doporučit robustní korelační koeficient založený na metodě nejmenších uškaných čtverců (LTS) nebo nejmenších vážených čtverců (LWS) pro praktické použití. Nicméně je výpočet těchto statistických odhadů velmi zdoluhavý, a to dokonce i při použití rychlých přibližných algoritmů.

Počáteční vyhlazení šumu bylo motivováno potřebou získat metodu robustní vůči šumu v obrazech. Efekt takového vyhlazení může být poměrně malý, obzvláště pokud používáme robustní míry korelace s vysokým bodem selháním jako míry podobnosti mezi obrazem a šablonou. Zatímco teoretické robustní vlastnosti zůstávají úkolem pro budoucí výzkum, metody popsáné v tomto článku

lze doporučit pro praktické použití také v molekulární genetice [17], [18] nebo v antropologickém výzkumu [3].

Na rozdíl od standardních postupů jsme pracovali s korelační analýzou při lokalizaci objektů v originálních (surových) obrazech, aniž bychom použili obvyklou redukci dimenze a extrakci příznaků. Standardní postupy používají normalizované pozice nebo invariantní deskriptory pro odstranění nebo potlačení efektu otočení obrazu nebo efektu změny velikosti obrazu. I tak tento článek vyvrací populární názor, že statistické metody nezvládnou úkol zpracovat originální obrazy. Slibné výsledky jsme získali s váženým Pearsonovým korelačním koeficientem a také s robustními verzemi Pearsonova korelačního koeficientu. Jako otevřený problém zde zůstává výpočet optimálních vah pro vážený korelační koeficient, které by zvýšily diskriminaci mezi těmi částmi obrazu, které odpovídají šabloně, a těmi, které jí neodpovídají. Takový problém je třeba řešit opět robustním způsobem. Abychom nezískali vysoce influenční váhy, je třeba regularizovat danou úlohu tak, že se zavede určitá horní mez na optimální váhy. Zajímavou úlohou by také bylo srovnání standardních klasifikačních metod mnohorozměrné statistické analýzy a jejich chování při lokalizaci orientačních bodů v dané databázi obrazů.

Poděkování

Tato práce vznikla v rámci Centra biomedicínské informatiky, projekt 1M06014 Ministerstva školství, mládeže a tělovýchovy České republiky. Autor děkuje dvěma anonymním recenzentům za hodnotné rady, které vedly k vylepšení článku.

Literatura

- [1] Yang M.-H., Kriegman D.J., Ahuja N. (2002): Detecting faces in images: A survey. *IEEE Trans. Pattern Anal. and Machine Intel.* 24, No. 1, 34-58.
- [2] Viola P., Jones M.J. (2004): Robust real-time face detection. *Int. Journal of Computer Vision* 57, 137-154.
- [3] Kalina J. (2010): Locating landmarks using templates. *Nonparametrics and Robustness in a Broader Perspectives. A Festschrift in Honor of Professor Jana Jurečková. IMS Collections No. 7.* Accepted, in print.
- [4] Dobeš M., Machala L., Tichavský P., Pospíšil J. (2004): Human eye iris recognition using the mutual information, *Optik* 115, 399-404.
- [5] Lin Z., Davis L.S., Doermann D.S., DeMenthon D. (2007): Hierarchical part-template matching for human detection and segmentation, in *Proceedings of the Eleventh IEEE International Conference on Computer Vision ICCV 2007*, IEEE Computer Society, Washington, 2007, 1-8.
- [6] Wolf L., Huang X., Martin I., Metaxas D. (2006): Patch-Based Texture Edges and Segmentation. In *Leonardis A., Bischof H., Pinz A. (Eds.): Computer Vision - ECCV 2006, 9th European Conference on Computer Vision, Graz, Proceedings, Part II. Lecture Notes in Computer Science 3952.*
- [7] Papageorgiou C.P., Oren M., Poggio T. (1998): A general framework for object detection, in *Proceedings of the Sixth IEEE International Conference on Computer Vision ICCV 1998*, IEEE Computer Society, Washington, 555-562.
- [8] Dalal N., Triggs B. (2005): Histograms of oriented gradients for human detection, in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR 2005*, IEEE Computer Society, Washington, 2005, pp. 886-893.
- [9] Torralba A., Murphy K.P., Freeman W.T. (2007): Sharing visual features for multiclass and multiview object detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 5, 854-869.
- [10] Tuzel O., Porikli F., Meer P. (2007): Human detection via classification on Riemannian manifolds, in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR 2007*, IEEE Computer Society, Washington.
- [11] Böhringer S., Vollmar T., Tasse C., Würtz R.P., Gillesen-Kaesbach G., Horsthemke B., Wiczorek D. (2006): Syndrome identification based on 2D analysis software. *Eur. J. Hum. Genet.* 14, 1082-1089.
- [12] Loos H.S., Wiczorek D., Würtz R.P., Malsburg von der C., Horsthemke B. (2003): Computer-based recognition of dysmorphic faces. *Eur. J. Hum. Genet.* 11, 555-560.

- [13] Würtz R.P. (1997): Object recognition robust under translations, deformations, and changes in background. *IEEE Trans. Pattern Anal. and Machine Intel.* 19, No. 7, 769-775.
- [14] Pitas I., Venetsanopoulos A.N. (1990): *Nonlinear digital filters*. Kluwer, Dordrecht.
- [15] Jurečková J., Pícek J. (2006): *Robust statistical methods with R*. Chapman & Hall/CRC, Boca Raton.
- [16] Arya K.V., Gupta P., Kalra P.K., Mitra P. (2007): Image registration using robust M-estimators, *Pattern Recognition Letters* 28, 1957-1968.
- [17] Dunning M.J., Smith M.L., Ritchie M.E., Tavaré S. (2007): beadarray: R classes and methods for Illumina bead-based data, *Bioinformatics* 23, 2183-2184.
- [18] Kalina J. (2010): Robust image analysis in the evaluation of gene expression studies. *ERCIM News, European Research Consortium for Informatics and Mathematics*, No. 82, p. 52.
- [19] Víšek J.Á. (2001): Regression with high breakdown point. In Antoch J., Dohnal G. (Eds.): *Proceedings of ROBUST 2000, Summer School of JČMF, JČMF and Czech statistical society*, 324-356.
- [20] Kalina J. (2007): Locating the mouth using weighted templates. *Journal of applied mathematics, statistics and informatics* 3, No. 1, 111-125.
- [21] Čížek P. (2008): Efficient robust estimation of time-series regression models. *Appl. Math.* 53, No. 3, 267-279.
- [22] Rousseeuw P.J., van Driessen K. (1999): A fast algorithm for the minimum covariance determinant estimator, *Technometrics* 41, 212-223.
- [23] Rousseeuw P.J., Leroy A.M. (1987): *Robust regression and outlier detection*. Wiley, New York.
- [24] Graf H.P., Cosatto E., Gibbon D., Kocheisen M., Petajan E. (1996): Multi-modal system for locating heads and faces. *Second IEEE International Conference on Automatic Face and Gesture Recognition FG 1996*, 88-93.
- [25] James M. (1987): *Pattern recognition*. BSP Professional books, Oxford.
- [26] Seshadri K., Savvides M. (2009): Robust modified active shape model for automatic facial landmark annotation of frontal faces, *Proceedings of the 3rd IEEE international conference on Biometrics: Theory, applications and systems*, IEEE Press, Piscataway, 319-326.
- [27] Shevlyakov G.L., Vilchevski N.O. (2001): *Robustness in data analysis: criteria and methods*. VSP, Utrecht.
- [28] Hampel F.R., Ronchetti E.M., Rousseeuw P.J., Stahel W.A. (1986): *Robust statistics, The approach based on influence functions*. Wiley, New York.

Kontakt

RNDr. Jan Kalina, Ph.D.,
 Ústav informatika AV ČR, v.v.i.
 Pod Vodárenskou věží 2
 182 07 Praha 8
 e-mail: kalina@euromise.cz

Spolehlivost složených dichotomních měření

Patricia Martinková¹, Karel Zvára²

¹Centrum biomedicínské informatiky AV ČR, Ústav informatiky AV ČR, v.v.i., Praha, Česká republika

²Oddělení medicínské infomatiky, Ústav informatiky AV ČR, v.v.i., Praha, Česká republika

Souhrn

Spolehlivost, neboli reliabilita měření je mírou jeho opakovatelnosti za stejných podmínek. Klasický koncept reliability předpokládá, že se měření Y skládá ze skutečné hodnoty měřené vlastnosti T a chybové složky ϵ , dvou nezávislých náhodných veličin, $Y = T + \epsilon$. Spolehlivost měření je pak definovaná jako podíl rozptylu skutečného skóru a rozptylu měření. Tento koncept nicméně není použitelný v modelech pro dichotomní měření, ve kterých nefigurují chybové složky, leč jsou definovány pomocí podmíněných pravděpodobností. V tomto článku zkoumáme obecnější definici spolehlivosti navrženou v [1], která je založena na rozkladu rozptylu v modelu se smíšenými efekty. Navržená definice splývá v klasické testové situaci s definicí klasickou, a navíc je použitelná také pro modely dichotomních měření. Nově jsou pro navrženou definici odvozeny předpoklady, za jejichž platnosti může být spolehlivost složeného měření vyjádřena pomocí spolehlivosti jediného měření (tzv. Spearmanova-Brownova formule) a je ukázána přibližná platnost Spearmanovy-Brownovy formule pro Raschův model. Na závěr zkoumáme tzv. *logistické alfa* - nový odhad spolehlivosti navržený v [2] coby modifikace klasického odhadu spolehlivosti založeném na Cronbachovu alfa. Simulace ukazují, že nový odhad nepodhodnocuje skutečnou spolehlivost tak často, jako Cronbachovo alfa. Logistické alfa se tak pro některé případy jeví být vhodnějším odhadem spolehlivosti. Nový odhad je použit na binární data získaná automatizovaným procesem diagnostiky prokrvení myokardu založeným na snímcích emisní protonové tomografie (SPECT).

Klíčová slova: spolehlivost měření, binární data, logistická regrese, Cronbachovo alfa, Raschův model, emisní protonová tomografie (SPECT)

1. Úvod a klasické pojetí spolehlivosti

Spolehlivost měření popisuje, do jaké míry bude výsledek stejný, zachováme-li stejné podmínky měření. V medicínské praxi spolehlivost měření stále zůstává důležitým tématem, které vzbuzuje mnoho diskuzí. Pro spojitá měření byla spolehlivost studována autory Yi, Wang a He. V jejich práci [3] byl také navržen test ekvivalence dvou měření. Neparametrický, pravděpodobnostní odhad spolehlivosti použitý na kognitivní testy pro pacienty s Alzheimerovou chorobou byl studován autory van Bellem a Arnoldovou [4]. Odhady reliability založené na IRT modelech, které jsou vhodné pro dichotomní nebo ordinální veličiny, byly použity v práci Teresiové a kol. [5].

Ve studiích zkoumajících opakovatelnost dichotomních měření jsou data obvykle analyzována s použitím kappa statistik, které byly motivovány vztahem ke koeficientu vnitrotřídní korelace [6], [7]. V této práci rozebíráme jiný přístup ke spolehlivosti: inspirování článkem [8] navrhujeme novou definici spolehlivosti založenou na rozkladu rozptylů v modelu se smíšenými efekty. Nová definice pokrývá klasickou testovou situaci a je navíc vhodná i pro modely dichotomních měření. Studujeme také nový odhad spolehlivosti. Pro lepší pochopení souvislosti s klasickou testovou teorií předkládáme v této části souhrn základních principů klasické testové teorie (CTT), z níž naše úvahy vychází.

1.1 Spolehlivost měření v rámci CTT

V klasické testové teorii [9] se předpokládá, že měření Y se skládá ze skutečné hodnoty měřené vlastnosti T a chybové složky ϵ , dvou nezávislých náhodných veličin

$$\begin{aligned} Y &= T + \epsilon, \\ T &\sim (\mu, \sigma_T^2), \sigma_T^2 > 0, \\ \epsilon &\sim (0, \sigma^2), \sigma^2 > 0. \end{aligned} \quad (1)$$

Spolehlivost měření je definovaná jako podíl rozptylu skutečného skóru a rozptylu naměřených hodnot

$$reli(Y) = \frac{\text{var}(T)}{\text{var}(Y)} = \frac{\sigma_T^2}{\sigma_T^2 + \sigma^2} = \rho_1. \quad (2)$$

Alternativně může být spolehlivost definována jako čtverec korelace naměřené hodnoty a skutečného skóru

$$\text{corr}^2(Y, T) = \text{corr}^2(T + \epsilon, T) = \frac{(\sigma_T^2)^2}{(\sigma_T^2 + \sigma^2)\sigma_T^2} = reli(Y). \quad (3)$$

Spolehlivost může být také vyjádřena jako korelace dvou opakovaných měření Y_1, Y_2 tedy dvou nezávislých, stejně přesných měření téže vlastnosti T

$$Y_j = T + \epsilon_j, j = 1, 2, \quad (4)$$

$$\text{corr}(Y_1, Y_2) = \text{corr}(T + \epsilon_1, T + \epsilon_2) = \frac{\sigma_T^2}{\sigma_T^2 + \sigma^2} = reli(Y).$$

Korelace mezi dvěma nezávislými (ne nutně stejně přesnými) měřeními téže vlastnosti T může být vyjádřena jako

$$Y_j = T + \epsilon_j, j = 1, 2,$$

$$\epsilon_1 \sim (0, \sigma_1^2), \epsilon_2 \sim (0, \sigma_2^2) \text{ nezávislé}$$

$$\begin{aligned} \text{corr}(Y_1, Y_2) &= \text{corr}(T + \epsilon_1, T + \epsilon_2) = \\ &= \frac{\sigma_T^2}{\sqrt{\sigma_T^2 + \sigma_1^2}\sqrt{\sigma_T^2 + \sigma_2^2}} = \sqrt{reli(Y_1)}\sqrt{reli(Y_2)} \end{aligned} \quad (5)$$

1.2 Spolehlivost součtu opakovaných měření

Máme-li J opakovaných měření vlastnosti T

$$Y_j = T + \epsilon_j, \quad \epsilon_j \sim (0, \sigma^2) \text{ iid}, \quad j = 1, \dots, J, \quad (6)$$

variabilita jejich součtu $Y_\bullet = \sum_j Y_j$ je

$$\text{var}(Y_{\bullet}) = \text{var}\left(\sum_j (T + \epsilon_j)\right) = J^2 \sigma_T^2 + J \sigma^2, \quad (7)$$

tudíž spolehlivost součtu p_j nebo průměru Y_{\bullet} můžeme vyjádřit pomocí spolehlivosti (každého) jednoho měření ρ_1 (tzv. Spearmanova-Brownova formule, viz [10], [11]).

$$\begin{aligned} \text{reli}(Y_{\bullet}) &= \frac{\text{var}(JT)}{\text{var}(Y_{\bullet})} = \frac{J^2 \sigma_T^2}{J^2 \sigma_T^2 + J \sigma^2} = \\ &= \frac{J \rho_1}{(J-1) \rho_1 + 1} = \rho_J. \end{aligned} \quad (8)$$

Spolehlivost součtu J opakovaných měření může být také vyjádřena pomocí kovariancí

$$\sigma_{jk} = \text{cov}(Y_j, Y_k)$$

jako

$$\begin{aligned} \text{reli}(Y_{\bullet}) &= \frac{\text{var}(JT)}{\text{var}(Y_{\bullet})} = \frac{J}{J-1} \frac{J-1}{J} \frac{J^2 \sigma_T^2}{\text{var}(Y_{\bullet})} = \\ &= \frac{J}{J-1} \frac{J^2 \sigma_T^2 + J \sigma^2 - J \sigma^2 - J \sigma_T^2}{\text{var}(Y_{\bullet})} \\ &= \frac{J}{J-1} \frac{\text{var}(\sum_j Y_j) - \sum_j \text{var}(Y_j)}{\text{var}(\sum_j Y_j)} = \\ &= \frac{J}{J-1} \frac{\sum_{j \neq k} \sigma_{jk}}{\sum_{j,k} \sigma_{jk}} = \alpha \end{aligned} \quad (9)$$

Ve vyjádření (9), jsme se dopracovali k tzv. **Cronbachovu alfa** (viz [12]), které lze snadno odhadnout, nahradíme-li populační kovariance jejich výběrovými protějšky

$$\hat{\alpha} = \frac{J}{J-1} \frac{\sum_{j \neq k} s_{jk}}{\sum_{j,k} s_{jk}},$$

$$\text{kde } s_{jk} = \frac{1}{I-1} \sum_{t=1}^I (Y_{tj} - \bar{Y}_{\bullet j})(Y_{tk} - \bar{Y}_{\bullet k}). \quad (10)$$

Pro dichotomní data odhad (10) splývá s Kuderovou-Richardsonovou formulí 20 (viz [13]).

Jak bylo výše ukázáno v (9), Cronbachovo alfa je ekvivalentní spolehlivosti součtu opakovaných měření **těže** vlastnosti T . Často je nicméně používáno pro odhadování spolehlivosti složených měření, viz dále.

1.3 Spolehlivost složených měření

Často nastávají situace, kdy měření nemohou být zopakována zcela nezávisle a měřit tak zcela stejnou vlastnost T . Například v psychometrii jsou testy složené z J položek, z nichž každá se soustředí na nepatrně jiný aspekt měřené vlastnosti. Každý subjekt je pak popsán součtem J položkových skóre $Y_{\bullet} = \sum Y_j$. V medicínské praxi se zase lékaři setkávají s měřeními od různých hodnotitelů nebo od téhož hodnotitele, který používá různé nástroje k měření. Měřená vlastnost je pak často popisována pomocí průměru takových měření $\bar{Y}_{\bullet} = \sum Y_j / J$.

Pro j -té měření předpokládáme, že

$$Y_j = T_j + \epsilon_j, \quad \epsilon_j \sim (0, \sigma_j), \quad j = 1, \dots, J \quad (11)$$

kde $(\epsilon_1, \dots, \epsilon_j)$ jsou vzájemně nezávislé a také nezávislé na (T_1, \dots, T_j) . Spolehlivost složeného měření pak je

$$\text{reli}(Y_{\bullet}) = \frac{\text{var}(T_{\bullet})}{\text{var}(T_{\bullet}) + \text{var}(\epsilon_{\bullet})}. \quad (12)$$

Jak bylo ukázáno Novickem a Lewisem v práci [14], Cronbachovo alfa je obecně dolní mezí spolehlivosti složeného měření

$$\alpha \leq \text{reli}(Y_{\bullet}) = \rho_J. \quad (13)$$

Novick a Lewis dále ukázali, že rovnost nastává právě v případě tzv. esenciálně τ -ekvivalentních položek, tedy v případě, kdy pro náhodnou veličinu T a nějaká reálná čísla β_j taková, že $\sum \beta_j = 0$, platí, že s pravděpodobností = 1

$$T_j = T + \beta_j, \quad \forall j = 1, \dots, J. \quad (14)$$

To je ekvivalentní požadavku, aby současně platily podmínky

$$\text{var}(T_j) = \sigma_T^2, \quad \forall j \quad (15)$$

$$\text{corr}(T_j, T_k) = 1, \quad \forall j, k. \quad (16)$$

Podmínka shody rozptylů (15) je zapotřebí, abychom dosáhli shodných spolehlivostí jednotlivých měření v modelu (11). Aby platila Spearmanova-Brownova formule, je pak také zapotřebí podmínka jedničkových korelačních koeficientů (16).

Nejsou-li položky složeného měření esenciálně τ -ekvivalentní, vedle skutečnosti, že neplatí Spearmanova-Brownova formule, můžeme také díky vztahu (13) očekávat, že odhad založený na Cronbachovu alfa bude podhodnocovat skutečnou spolehlivost složeného měření. Odhady neshody mezi spolehlivostí a Cronbachovým alfa na populační úrovni již byly studovány v práci [15].

1.4 Cronbachovo alfa jako odhad spolehlivosti složeného měření

Předpokládejme, že máme měření Y_{ij} , $j = 1, \dots, J$ na subjektech. Předpoklad esenciální τ -ekvivalence vede ke smíšenému modelu analýzy rozptylu dvojnásobného třídění.

$$Y_{ij} = T_i + \beta_j + \epsilon_{ij}, \quad \sum_j \beta_j = 0, \quad i = 1, \dots, I, \quad j = 1, \dots, J \quad (17)$$

kde β_j je parametr měření (obtížnost položky, úroveň daného hodnotitele, atd.). Přidáme-li požadavky na normalitu

$$T_i \sim N(\mu, \sigma_T^2), \quad \epsilon_{ij} \sim N(0, \sigma^2),$$

a vezmeme-li v potaz rozklad součtu

$$SS_T = \sum (Y_{ij} - \bar{Y}_{\bullet\bullet})^2 = SS_A + SS_B + SS_E,$$

čtverců pak průměrné čtverce MS_A a MS_E mají následující střední hodnoty:

$$E M S_A = E \sum_{i=1}^I \sum_{j=1}^J (\bar{Y}_{i\bullet} - Y_{i\bullet})^2 / (I-1) = J \sigma_T^2 + \sigma^2, \quad (18)$$

$$E M S_E = E \sum_{i=1}^I \sum_{j=1}^J (Y_{ij} - \bar{Y}_{i\bullet} - \bar{Y}_{\bullet j} + \bar{Y}_{\bullet\bullet})^2 / ((I-1)(J-1)) = \sigma^2. \quad (19)$$

Tudíž můžeme Cronbachovo alfa (9) vyjádřit jako

$$\alpha = \frac{E M S_A - E M S_E}{E M S_A} \quad (20)$$

a odhad (10) může být přepsán jako

$$\hat{\alpha} = \frac{M S_A - M S_E}{M S_A} = 1 - \frac{M S_E}{M S_A} = 1 - \frac{1}{F_T}, \quad (21)$$

kde F_T je statistika běžně používaná pro testování podmodelu bez efektu jedince v plném modelu (17). Stejná statistika se používá také pro testování podmodelu v modelu analýzy rozptylu dvojitěho třídění s pevnými efekty [16].

Ze vztahu (21) můžeme odvodit, že $\hat{\alpha}$ nabývá hodnot blízkých jedničce v případě, že složené měření dobře rozlišuje mezi jedinci. Tudíž můžeme také očekávat, že v případě větší variability mezi jedinci bude také $\hat{\alpha}$ dosahovat vyšších hodnot.

2. Původní výsledky pro dichotomní měření

V případě, kdy měření Y nabývají pouze hodnoty 0 nebo 1, není klasický model popsán v části 1.1 vyhovující, neboť Y nelze vyjádřit jako součet dvou nezávislých náhodných veličin. Model v takovém případě definujeme spíše pomocí podmíněných středních hodnot $E(Y|T)$. Předpokládejme proto obecnější model

$$Y_{ij} \sim f(\cdot, T_i), \quad T_i \sim iid,$$

$$(Y_{ij}|T_i), (Y_{i'j'}|T_{i'}) \text{ nezávislé pro } j \neq j' \quad (22)$$

Jedním z takových modelů je také Raschův model (viz [17]):

$$E(Y_{ij}|T_i) = \pi(T_i, \beta_j) = \frac{\exp(T_i + \beta_j)}{1 + \exp(T_i + \beta_j)}, \quad T_i \sim N(\mu, \sigma_T^2). \quad (23)$$

V rámci tzv. Item Response Theory (IRT) jsou právě Raschův model (23) a mnohá jeho další zobecnění obsáhle studována [18], a to zvláště v souvislosti s problematikou odhadu parametrů modelu. Koncept spolehlivosti je v rámci IRT rozšířen z jediného indexu na funkci skutečného skóru T která se zde nazývá *testová informační funkce* [19]. Nicméně je zde také možné obdržet index pro test jako celek, který je přímou analogií Cronbachova alfa: jako analogie k rozložení pozorovaného měření na pravou a chybovou složku v klasické testové teorii, zde uvažujeme rozklad odhadu skutečného skóru na skutečný skór a chybovou složku

$$\hat{T} = T + \epsilon. \quad (24)$$

Spolehlivost v IRT je pak definována jako

$$R_{IRT} = \frac{\text{var}(T)}{\text{var}(\hat{T})} = \frac{\text{var}(\hat{T}) - \text{var}(\epsilon)}{\text{var}(\hat{T})}. \quad (25)$$

Pro odhadování této IRT spolehlivosti se nejprve odhadnou úrovně skutečných skóru jednotlivých subjektů a pomocí standardních procedur jsou také spočítány jejich střední chyby (SE). Poté je spočten výběrový rozptyl těchto odhadů pro odhad rozptylu $\text{var}(\hat{T})$. Odhadnutá střední chyba je základem pro odhad rozptylu chyby $\text{var}(\epsilon)$. Výsledný odhad bývá velice blízký Cronbachovu alfa [20].

V tomto článku uvažujeme novou definici spolehlivosti pro model (22). Jedná se o přímočařejší zobecnění klasické definice spolehlivosti než je (25). Uvádíme také vzorec pro spolehlivost v Raschově modelu a předpoklady pro platnost Spearmanovy-Brownovy formule v modelu (22). Diskutujeme nový odhad spolehlivosti, vhodný pro složená

dichotomní měření a porovnááme jej s Cronbachovým alfa v simulacích a v praktickém příkladě.

2.1 Navrhovaná definice spolehlivosti
Rozptyl měření Y_{ij} řídicího se modelem (22) může být rozložen pomocí podmíněného rozptylu a podmíněné střední hodnoty na složky

$$\text{var}(Y_{ij}) = E(\text{var}(Y_{ij}|T_i)) + \text{var}(E(Y_{ij}|T_i)). \quad (26)$$

První člen je vnitřní rozptyl, tedy ta část rozptylu, která nesouvisí s variabilitou skutečného skóru T . Druhý člen je rozptyl mezitřídí, tedy ta část rozptylu, která odpovídá variabilitě skutečného skóru T [8].

S odkazem na klasickou definici spolehlivosti můžeme nyní definovat spolehlivost jako podíl rozptylu odpovídajícího variabilitě skutečného skóru k celkovému rozptylu naměřených hodnot, tedy jako

$$\text{reli}(Y) = \frac{\text{var}(E(Y|T))}{\text{var}(Y)} = \frac{\text{var}(E(Y|T))}{\text{var}(E(Y|T)) + E(\text{var}(Y|T))}. \quad (27)$$

Jelikož pro klasický model platí

$$E(Y|T) = E((T + E)|T) = T,$$

definice (27) v tomto případě splývá s klasickou definicí spolehlivosti (2).

Pro složená měření definujeme spolehlivost j -té položky jako

$$\text{reli}(Y_{ij}) = \frac{\text{var}(E(Y_{ij}|T_i))}{\text{var}(Y_{ij})} = \tau_{ij} \quad (28)$$

a spolehlivost složeného měření $Y_{i\bullet}$ může být definována jako

$$\text{reli}(Y_{i\bullet}) = \frac{\text{var}(E(Y_{i\bullet}|T_i))}{\text{var}(Y_{i\bullet})}. \quad (29)$$

V práci [1] jsme odvodili, že spolehlivost složeného měření v Raschovu modelu (23) lze vyjádřit jako

$$\text{reli}(Y_{\bullet}) = \frac{\text{var}(E(Y_{\bullet}|T))}{\text{var}(Y_{\bullet})} = \frac{\sum_{j=1}^J \sum_{t=1}^J (C_{jt} - D_j D_t)}{\sum_{j=1}^J \sum_{t=1}^J (C_{jt} - D_j D_t) + \sum_{j=1}^J B_j}, \quad (30)$$

kde integrály

$$B_j = E_T \frac{e^{T+\beta_j}}{(1 + e^{T+\beta_j})^2},$$

$$D_j = E_T \frac{e^{T+\beta_j}}{1 + e^{T+\beta_j}},$$

$$C_{jt} = E_T \frac{e^{T+\beta_j}}{1 + e^{T+\beta_j}} \frac{e^{T+\beta_t}}{1 + e^{T+\beta_t}}$$

neumíme vyjádřit explicitně, leč mohou být vyčísleny numerickými metodami. Pro danou testovou situaci (pro rozdělení skutečných úrovní subjektů T , počet položek J a jejich úrovní β_j , $j = 1, \dots, J$) tak umíme vyčíslit skutečnou hodnotu spolehlivosti (viz Tabulka 1).

2.2 Spearmanova-Brownova formule

Pokusme se nejprve najít podmínky, za kterých si budou spolehlivosti jednotlivých měření τ_j rovny pro všechna j . Jak jsme již dříve zmínili, spolehlivost můžeme vyjádřit jako korelaci dvou nezávislých měření téže vlastnosti T (viz vztah (4)), tedy pomocí vnitrotřídní korelace (náležící i -tému subjektu) $\rho_{ijj'} = \text{corr}(Y_{ij} Y_{ij'})$.

V práci [21] jsme ukázali, že v rámci modelu (22) je vztah mezi ρ_{ij} a τ_j pro $j \neq j'$ následující:

$$\rho_{ijj'} = \sqrt{\tau_{ij}} \sqrt{\tau_{ij'}} \text{corr}[E(Y_{ij}|T_i)],$$

$$E(Y_{ij'}|T_i) \leq \sqrt{\tau_{ij}} \sqrt{\tau_{ij'}}. \quad (31)$$

Rovnost ve vztahu (31) platí pro všechna $j \neq j'$ jestliže pro všechna $j \neq j'$

$$\text{corr}[E(Y_{ij}|T_i), E(Y_{ij'}|T_i)] = 1, \quad (32)$$

což je právě tehdy, když pro všechna i, j s jedničkovou pravděpodobností pro nějaké konstanty $k_{ij} > 0$ a η_{ij} a pro nějaké funkce $\lambda_i(T)$ mohou být podmíněné střední hodnoty vyjádřeny jako

$$E(Y_{ij}|T_i) = k_{ij} [\lambda_i(T_i) + \eta_{ij}]. \quad (33)$$

Navíc můžeme k podmínce (33) požadovat splnění dodatečné podmínky

$$\sum_{j=1}^J k_{ij}^2 = J.$$

jelikož λ_i a η_{ij} lze vždy vynásobit vhodnými konstantami.

Formule (31) a předpoklad (33) mohou být rozšířeny do následujícího tvrzení (pro důkaz viz [21]), které reviduje tvrzení autorů Commengese a Jacqminové (viz [8]):

Tvrzení Předpokládejme, že se Y_{ij} řídí pro $i = 1, \dots, I$, $j = 1, \dots, J$, $J \geq 3$ modelem (22). Navíc necht' s jedničkovou pravděpodobností platí pro všechna i, j podmínka (33). Pak jsou následující tři tvrzení ekvivalentní:

P1 $\rho_{ij} = \rho_i$, nezávisí na j, j' pro žádné $j \neq j'$.

P2 $\tau_j = \tau_i$, nezávisí na j pro žádné j' .

P3 Model náleží do třídy modelů specifikované (s pravděpodobností rovnou jedné) vztahem:

$$\text{var}(Y_{ij}|T_i) = k_{ij}^2 [\sigma_i^2(T_i) + \psi_{ij}(T_i)],$$

$$\sum_{k=1}^J k_{ij}^2 = J, \quad (34)$$

kde $E[\psi_{ij}(T_i)] = 0$, a $\psi_{ij}(T_i) > -\sigma_i^2(T_i)$.

Navíc tato tři tvrzení implikují, že

P4 $\rho_i = \tau_i$.

Odtud plyne, že platí-li podmínky (33) a (34), pak jsou spolehlivosti všech jednotlivých měření j rovny τ_i . Podívejme se nyní na spolehlivost složeného měření za platnosti těchto dvou podmínek a za platnosti zmíněné dodatečné podmínky

$$\sum_{j=1}^J k_{ij}^2 = J.$$

Spolehlivost (každého) jednoho měření můžeme vyjádřit jako

$$R_1 = \tau_{ij} = \frac{\text{var}[E(Y_{ij}|T_i)]}{\text{var}[E(Y_{ij}|T_i)] + E[\text{var}(Y_{ij}|T_i)]} =$$

$$= \frac{\text{var}[\lambda_i(T_i)]}{\text{var}[\lambda_i(T_i)] + E[\sigma_i^2(T_i)]} = \tau_i = \rho$$

a ta dle Tvrzení splývá s korelací mezi dvěma měřeními téže vlastnosti Y_{ij} , $Y_{ij'}$, $j \neq j'$. Spolehlivost složeného měření lze pak vyjádřit jako

$$R_J = \frac{\text{var}[E(Y_i|T_i)]}{\text{var}[E(Y_i|T_i)] + E[\text{var}(Y_i|T_i)]} =$$

$$= \frac{\text{var}[\sum_{j=1}^J E(Y_{ij}|T_i)]}{\text{var}[\sum_{j=1}^J E(Y_{ij}|T_i)] + \sum_{j=1}^J E[\text{var}(Y_{ij}|T_i)]}$$

$$= \frac{\text{var}\{\sum_{j=1}^J k_{ij}[\lambda_i(T_i) + \eta_{ij}]\}}{\text{var}\{\sum_{j=1}^J k_{ij}[\lambda_i(T_i) + \eta_{ij}]\} + \sum_{j=1}^J k_{ij}^2 E[\sigma_i^2(T_i) + \psi_{ij}(T_i)]}$$

$$= \frac{(\sum_{j=1}^J k_{ij})^2 \text{var}[\lambda_i(T_i)]}{(\sum_{j=1}^J k_{ij})^2 \text{var}[\lambda_i(T_i)] + m E[\sigma_i^2(T_i)]} =$$

$$= \frac{(\sum_{j=1}^J k_{ij})^2 R_1}{1 + \left(\frac{\sum_{j=1}^J k_{ij})^2}{(\sum_{j=1}^J k_{ij})^2} - 1\right) R_1}. \quad (35)$$

Výraz (35) splývá se Spearmanovou-Brownovou formulí v případě, kdy $k_{ij} = 1$ pro všechna j . Dospíváme tedy k závěru, že podmínky esenciální τ -ekvivalence (15)(16) pro klasický model korespondují v modelu (22) s podmínkou, že s pravděpodobností rovnou jedné mohou být podmíněné střední hodnoty a podmíněné rozptyly měření Y_{ij} vyjádřeny jako

$$E(Y_{ij}|T_i) = \lambda_i(T_i) + \eta_{ij}, \quad (36)$$

$$\text{var}(Y_{ij}|T_i) = \sigma_i^2(T_i) + \psi_{ij}(T_i), \quad (37)$$

kde η_j jsou dané konstanty, $E[\psi_{ij}(T_i)] = 0$,
a $\psi_{ij}(T_i) > -\sigma_i^2(T_i)$.

Pojďme nyní prozkoumat platnost Spearmanovy-Brownovy formule v Raschovu modelu. Jak bylo ukázáno v [21], Raschův model (23) nespĺňuje podmínku (36) ani podmínku (37). Zkusme nicméně Spearmanovu-Brownovu formuli použít. V Tabulce 1 jsou spočteny hodnoty spolehlivosti v Raschovu modelu dle vzorce (30) pro různé testové situace (počet položek J , úrovně položek rozložené ekvidistantně na intervalu $\langle -0.1, 0.1 \rangle$ úrovně jednotlivých subjektů T_i se řídí normálním rozdělením s nulovou střední hodnotou a rozptylem σ_T^2). Integrály byly počítány pomocí funkce integrate v prostředí R (viz [22]), maximální absolutní chyba dosažená při integrování byla menší než 0.000025. V Tabulce 2 jsme jako pevný vzali počet položek $J = 11$ a použili jsme druhý řádek Tabulky 1 spolu se Spearmanovou-Brownovou formulí (8) k výpočtu přibližných hodnot spolehlivosti pro počty položek $J = 3, 20, 50$ a 100.

Jak můžeme vidět, hodnoty v Tabulkách 1 a 2 jsou velmi podobné a zdá se, že tedy Spearmanova-Brownova formule (8) pro Raschův model platí alespoň přibližně. Přibližnou platnost Spearmanovy-Brownovy formule můžeme prokázat s použitím Taylorova rozvoje: Předpokládejme malá b_j taková, že

$$\sum_j b_j = 0,$$

a použijme Taylorův rozvoj prvního stupně (na funkci jedné resp. dvou proměnných)

$$B_j = E \frac{e^{T+b_j}}{(1+e^{T+b_j})^2} \approx E \frac{e^T}{(1+e^T)^2} +$$

$$+ b_j E \frac{e^T(1-e^T)}{(1+e^T)^3} = B + b_j E \frac{e^T(1-e^T)}{(1+e^T)^3},$$

$$D_j = E \frac{e^{T+b_j}}{1+e^{T+b_j}} \approx E \frac{e^T}{1+e^T} +$$

$$+ b_j E \frac{e^T}{(1+e^T)^2} = D + b_j E \frac{e^T}{(1+e^T)^2},$$

Tab. 1. Spolehlivost v Raschovu modelu pro různé počty položek.

Počet položek	Rozptyl úrovně subjektů σ_T^2						
	0.01	0.1	0.2	0.5	0.9	2.5	10
J=3	0.00008	0.00741	0.02881	0.15047	0.34335	0.73121	0.94152
J=11	0.00028	0.02667	0.09814	0.39386	0.65731	0.90890	0.98335
J=20	0.00050	0.04747	0.16519	0.54160	0.77717	0.94775	0.99077
J=50	0.00125	0.11078	0.33098	0.74709	0.89711	0.97843	0.99629
J=100	0.00249	0.19947	0.49735	0.85524	0.94577	0.98910	0.99814

Tab. 2. Spearmanova-Brownova formule použitá pro $J=11$.

Počet položek	Rozptyl úrovně subjektů σ_T^2						
	0.01	0.1	0.2	0.5	0.9	2.5	10
SB R_3	0.00008	0.00742	0.02882	0.15054	0.34345	0.73125	0.94153
SB R_{20}	0.00050	0.04746	0.16518	0.54159	0.77716	0.94775	0.99077
SB R_{50}	0.00125	0.11077	0.33095	0.74707	0.89710	0.97843	0.99629
SB R_{100}	0.00249	0.19944	0.49731	0.85522	0.94576	0.98910	0.99814

$$C_{jt} = E \frac{e^{T+b_j}}{1+e^{T+b_j}} \frac{e^{T+b_t}}{1+e^{T+b_t}} \approx$$

$$\approx E \frac{e^{2T}}{(1+e^T)^2} + (b_j + b_t) E \frac{e^{2T}}{(1+e^T)^3} =$$

$$= C + (b_j + b_t) E \frac{e^{2T}}{(1+e^T)^3}.$$

Pak tedy spolehlivost složeného měření v Raschovu modelu lze přibližně vyjádřit jako

$$R_J \approx \frac{J^2(C-D^2)}{J^2(C-D^2) + JB}.$$

$$\hat{\alpha}_{log} = 1 - \frac{I-1}{X^2}. \quad (38)$$

Zatímco spolehlivost jednotlivých měření je přibližně

$$R_1 \approx \frac{C-D^2}{C-D^2+B},$$

což dohromady dává přibližnou platnost Spearmanovy-Brownovy formule v Raschovu modelu.

2.3 Logistické alfa odhad spolehlivosti složených dichotomních měření

F_T statistika v odhadu založeném na Cronbachovu alfa (21) nejlépe vyhovuje v případě normálně rozdělených dat. Pro dichotomní data bychom mohli zkusit zaměnit F_T analogickou statistikou z logistické regrese. V modelu s pevnými efekty logistické regrese je takovou vhodnou statistikou rozdíl deviancí podmodelu a modelu $X^2 = D(B) - D(A+B)$. Tato statistika má za nulové hypotézy asymptoticky (pro I pevné a J jdoucí do nekonečna) χ^2 rozdělení s $(I-1)$ stupni volnosti. Navržená analogie odhadu spolehlivosti, vhodná pro složená dichotomní měření, tzv. *logistické alfa* [2], [1] má tedy tvar:

V následující části pomocí simulací porovnáme nový odhad spolehlivosti, logistické alfa, s klasickým odhadem založeným na Cronbachovu alfa.

3. Simulační příklad

Tato simulace odpovídá následujícímu příkladu z lékařské praxe: $I = 20$ pacientů zodpovídalo na položek dotazníku kvality života.

Na každou položku bylo možné odpovědět pouze ano či ne. Úrovně jednotlivých položek β_j byly rozděleny ekvidistantně na intervalu $\langle -0.1, 0.1 \rangle$ úrovně skutečných skóre kvality života jednotlivých pacientů se řídily normálním rozdělením se střední hodnotou $\mu = 0$ a rozptylem σ_T^2 (55 hodnot rozptylů bylo vybráno z intervalu $\langle 0.01, 10 \rangle$ tak, aby odpovídajících 55 hodnot spolehlivosti složeného měření v Raschovu modelu daných vztahem (30) bylo přibližně rovnoměrně rozděleno na intervalu $(0, 1)$).

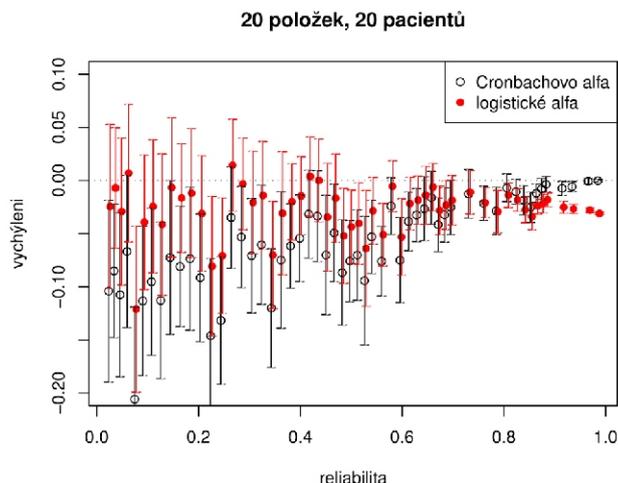
Pro každou kombinaci I, J a σ_T^2 byla pomocí formule (30) spočítána hodnota skutečné spolehlivosti složeného měření a bylo generováno 500 skupin dat: Z normálního rozdělení $N(0, \sigma_T^2)$ bylo generováno I úrovní T_i kvality života pacientů. Následně byly vygenerovány ano/ne odpovědi na dotazník kvality života z Raschova modelu (23). Dále byly na základě dat spočteny odhady $\hat{\alpha}_{CR}$ a $\hat{\alpha}_{log}$.

Ze získaných 500 odhadů $\hat{\alpha}_{CR}$ a 500 odhadů $\hat{\alpha}_{log}$ bylo spočítáno vychýlení spolu s jeho konfidenčním intervalem a dále střední čtvercová chyba, ty pak byly vykresleny do Obrázku 1 a 2.

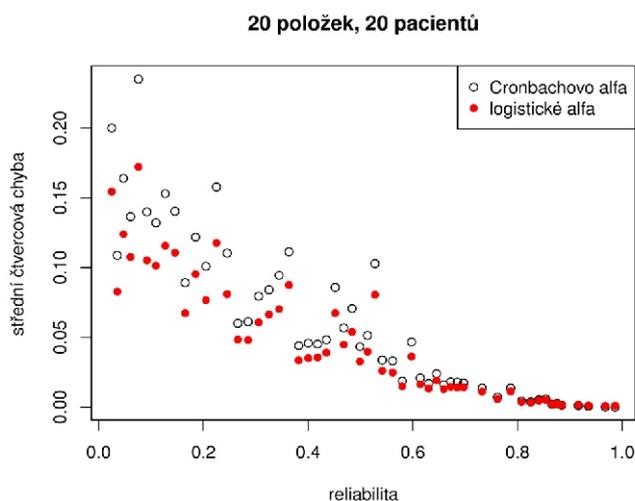
Byly zkoumány také další testové situace (počet položek $J = 11$ a 50 počet pacientů $I = 30$ a 50). Pozorovali jsme menší odhadnuté vychýlení a MSE pro logistické alfa $\hat{\alpha}_{log}$ obzvláště pro hodnoty skutečné spolehlivosti $\leq 0,75$. Poněkud horší výsledky *logistického alfa* byly získány pro spolehlivost blízkou 1 a pro vysoký počet pacientů v porovnání s počtem položek. Toto může být důsledkem skutečnosti, že zatímco statistika χ^2 použitá v (38) je vhodná pro model logistické regrese s pevnými efekty, v (23) předpokládáme model se smíšenými efekty. Šance na zlepšení odhadu by tedy mohla plynout z nahrazení statistiky χ^2 jinou, vhodnější statistikou.

4. Analýza kardiologických dat

Datový soubor **SPECT heart data** [23] popisuje diagnostikování průtoku myokardu na základě snímků protonové emisní tomografie (SPECT). Každý z 267 pacientů byl klasifikován do jedné ze dvou kategorií: normální a abnormální. Na



Obr. 1. Odhadnuté vychýlení pro klasický a logistický odhad spolehlivosti.



Obr. 2. Odhadnuté MSE pro klasický a logistický odhad spolehlivosti.

základě snímků protonové emisní tomografie bylo automatizovaně vygenerováno 22 binárních proměnných (22 dílčích diagnóz založených na SPECT snímcích). Pro detailnější popis procesu viz [24].

Nás zajímala vnitřní konzistence zmíněných 22 dílčích diagnóz založených na SPECT. Dosažené odhady v tomto případě byly: Cronbachovo alfa $\hat{\alpha}_{CR} = 0.839$ a logistické alfa $\hat{\alpha}_{log} = 0.827$. V tomto případě daly oba odhady velmi podobné hodnoty a ukazují na poměrně vysokou vnitřní konzistenci dílčích diagnóz.

5. Závěry a diskuze

V tomto článku jsme shrnuli základní principy klasické testové teorie (CTT) a použili jsme je pro objasnění motivace pro naši navrženou novou, obecnější definici spolehlivosti a nového odhadu spolehlivosti, vhodného pro složená dichotomní měření. Ukázali jsme, že navržená definice v klasické testové situaci splyvá s definicí spolehlivosti zavedenou v CTT. Odvodili jsme podmínky pro platnost Spearmanovy-Brownovy formule v modelu (22), který je obecnější než klasický model (1).

Navržená definice spolehlivosti byla aplikována na Raschův model. Pro ten bylo ukázáno, že v něm Spearmanova-Brownova formule neplatí zcela, ale platí alespoň přibližně. O navrženém odhadu spolehlivosti složených dichotomních měření, tzv. *logistickém alfa*, bylo na základě simulací ukázáno, že v jistých případech dává lepší výsledky (nižší odhadnuté vychýlení a MSE), obzvláště pro spolehlivost $\leq 0,75$ a pro počet položek převyšující počet subjektů. Byly diskutovány možnosti vylepšení odhadu spolehlivosti také pro spolehlivost blízkou a pro vyšší počet subjektů. Odhad spolehlivosti složených dichotomních měření byl demonstrován na datech z kardiologie.

Práce prezentovaná v tomto článku by mohla vést ke zpřesnění odhadování spolehlivosti složených dichotomních měření a přispět tak k mnoha oborům biomedicínského výzkumu.

Poděkování

Článek vznikl za podpory projektu 1M06014 Ministerstva školství, mládeže a tělovýchovy České republiky a AV0Z10300504 ÚI AV ČR. Autoři děkují recenzentům za jejich podnětné připomínky.

Literatura

- [1] Martinková P, Zvára K. Reliability in the Rasch model. *Kybernetika* 2007; 43(3):315326.
 [2] Zvára K. Measuring of reliability: Beware of Cronbach. [Měření reliability aneb bacha na Cronbacha, in Czech]. *Information Bulletin of the Czech Statistical Society* 2002; 12:1320.

- [3] Yi Q, Wang PP, He Y. Reliability analysis for continuous measurements: Equivalence test for agreement. *Statistics in Medicine* 2008; 27:28162825.
 [4] vanBelle G, Arnold A. Reliability of cognitive tests used in Alzheimer's disease. *Statistics in Medicine* 2000; 19:14111420.
 [5] Teresi JA, Kleinman M, Ocepek-Welikson K. Modern psychometric methods for detection of differential item functioning: application to cognitive assessment measures. *Statistics in Medicine* 2000; 19:16511683.
 [6] Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educational and Psychological Measurement* 1973; 33: 613619.
 [7] Kraemer HC. Ramifications of a population-model for kappa as a coefficient of reliability. *Psychometrika* 1979; 44(4):461472.
 [8] Commenges D, Jacqmin H. The intraclass correlation coefficient distribution-free definition and test. *Biometrics* 1994; 50:517526.
 [9] Suen HK. *Principles of Test Theories*. LEA Publishers, Hillsdale, New Jersey, 1990.
 [10] Spearman C. Correlation calculated from faulty data. *British Journal of Psychology* 1910; 3:271296.
 [11] Brown W. Some experimental results in the correlation of mental abilities. *British Journal of Psychology* 1910; 3:296322.
 [12] Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951; 16:297334.
 [13] Kuder G, Richardson M. The theory of estimation of test reliability. *Psychometrika* 1937; 2:151160.
 [14] Novick MR, Lewis C. Coefficient alpha and the reliability of composite measurement. *Psychometrika* 1967; 32:113.
 [15] Raykov T. Scale reliability, Cronbach's coefficient Alpha, and violations of essential tau-equivalence with fixed congeneric components. *Multivariate Behavioral Research* 1997; 32(4):329353.
 [16] Neter J, Wasserman W, Kutner MH. *Applied Linear Statistical Models*. Richard D. Irwin, Inc., Homewood, IL, USA, 1985
 [17] Rasch G. *Probabilistic Models for Some Intelligence and Attainment Tests*. The Danish Institute of Educational Research, 1960.
 [18] van der Linden WJ, Hambleton RK (editors). *Handbook of Item Response Theory*. Springer-Verlag, New York, 1997.
 [19] Samejima F. *Estimation of Reliability Coefficients Using the Test Information Function and Its Modifications*. *Applied Psychological Measurement* 1994; 18(3):229244.
 [20] Andrich D. An index of person separation in latent trait theory, the traditional KR-20 index, and the Guttman scale response pattern. *Educational Research and Perspectives* 1982; 9(1):95104.
 [21] Martinková P. *Reliability of Measurements Consisting of Dichotomously Scored Items*. Unpublished dissertation. Charles University, Prague, 2007.
 [22] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2010. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
 [23] Cios KJ, Kurgan LA. *SPECT heart data*. Colorado, USA, 2010. URL <http://archive.ics.uci.edu/ml/datasets/SPECT+Heart>.
 [24] Kurgan LA et al. Knowledge discovery approach to automated cardiac SPECT diagnosis. *Artificial Intelligence in Medicine* 2001; 23: 149-169.

Kontakt

Mgr. Patricia Martinková

RNDr. Patricia Martinková
 Ústav informatiky AV ČR, v.v.i,
 Centrum biomedicínské informatiky
 Pod Vodárenskou věží 2
 182 07 Praha 8
 e-mail: martinkova@euromise.cz



EJBI 2010 ISSN 1801 - 5603

European Journal for Biomedical Informatics

Volume 6 (2010), Issue 2

Croatian version

www.ejbi.eu



Content

Croatian version

- hr 2 - 4 [Akcijski tim za podizanje učinkovitosti EFMI-evih časopisa: Osnova, logika i svrha](#)
Izet Masic, Josipa Kern, Jana Zvárová, Simon de Lusignan, Gaj Vidmar

Akcijski tim za podizanje učinkovitosti EFMI-evih časopisa: Osnova, logika i svrha

Izet Masic¹, Josipa Kern², Jana Zvárová³, Simon de Lusignan⁴

¹Society for Medical Informatics of Bosnia and Herzegovina, Sarajevo, Bosnia and Herzegovina,

²Croatian Society for Medical Informatics, Zagreb, Croatia,

³Czech Society for Biomedical Engineering and Medical Informatics, Prague, Czech Republic,

⁴Division of Community Health Sciences, St George's University of London, UK,

Reprinted from: *Medical Informatics in a United and Healthy Europe*, K.P. Adlassnig et al. (Eds), IOS Press, 2009, 946-950, © 2009 European Federation for Medical Informatics.

Sažetak

Članak analizira način na koji EFMI širi nova znanja kao i postojeće medicinskoinformatičke časopise u zemljama članicama EFMI-a u skladu sa zaključkom sa sastanka EFMI-evog vijeća u Londonu 2008.godine. Identificirano je osam glavnih medicinskoinformatičkih časopisa i nekoliko drugih publikacija. Većina njih se zasniva na pretplati i izlazi barem kvartalno. Realna je mogućnost da urednici ustanove praksu redovitog sastajanja sa svrhom podizanja učinkovitosti u širenju novih znanja i dobre prakse u medicinskoj informatici. Moguća je i razmjena stručnih znanja i harmonizacija u pripremanju i slanju rukopisa kako bi se što više reducirale barijere pri objavljivanju radova.

Ključne riječi: EFMI, misija, akcijski tim za podizanje učinkovitosti medicinskoinformatičkih časopisa

1. Uvod

Europska federacija za medicinsku informatiku (EFMI), neprofitna znanstvena i stručna organizacija ustanovljena 1976. godine danas ima 32 zemlje članice (1-4). Sve europske zemlje imaju pravo da u EFMI-u učlane svoja medicinskoinformatička društva. Termin "medicinska informatika" uključuje cijeli spektar zdravstvene informatike i svih njezinih potpodručja. MeSH (engl. Medical Subject Headings) definira medicinsku informatiku kao "područje informacijskih znanosti koje se bavi analizom i diseminacijom medicinskih podataka primjenjujući računala u raznim aspektima zdravstvene zaštite i medicine" (4,5). Ova je definicija uvedena 1987. godine kao pojam u MeSH-u, rječniku međunarodno

usvojenih medicinskih pojmova.

EFMI pruža minimalnu administrativnu potporu nacionalnim društvima na način da predstavnici svih društava sudjeluju u funkcioniranju EFMI-evog vijeća. Službeni jezik je engleski, premda se katkada u zemljama u kojima engleski nije službeni jezik osigurava simultano prevođenje (6). EFMI ima 16 radnih grupa: Pametne kartice u zdravstvu „CARDS“; Društvena sigurnost i socijalna skrb; Edukacija u medicinskoj informatici „EDU“; Elektronički zdravstveni zapis „EHR“; Evaluacija zdravstvenih informacijskih sustava „EVAL“; Informatika za invalidne osobe i rehabilitaciju „IDR“; Planiranje i modeliranje informacija u zdravstvu „IPAM“; Besplatni i otvoreni izvorni programski kod „LIFOSS“; Mješoviti i ozbiljni slučajevi; Medicinska informatika u tranzicijskim zemljama „MCIT“; Obrada medicinskih slika „MIP“; Razumijevanje prirodnih jezika „NLU“; Informatika u sestrinstvu u Europi „NURSIE“; Ljudski i organizacijski čimbenici u medicinskoj informatici „HOFMI“; Informatika u primarnoj zdravstvenoj zaštiti „PCI“; Sigurnost, zaštita i etika „SSE“; Sljedivost „TRA“.

EFMI je dosad organizirao 21 europski kongres medicinske informatike pod nazivom Medical Informatics Europe (MIE): Cambridge (1978), Berlin (1979), Oslo (1988), Glasgow (1990), Beč (1991), Jerusalem (1993), Lisabon (1994), Copenhagen (1996), Thessaloniki (1997), Ljubljana (1999), Hannover (2000), Budimpešta (2002), St. Malo (2003), Ženeva (2005), Maastricht (2006), Gothenburg (2008). MIE 2009 održat će se u Sarajevu a MIE 2011 u Oslu. Tradicionalno, kongresi MIE se ne

održavaju u godinama u kojima Međunarodna asocijacija za medicinsku informatiku organizira svoje trienalne konferencije MEDINFO (1). Nedavno je EFMI započeo s organizacijom nove serije sastanaka, tzv. konferencije s posebnim temama (engl. Special Topic Conference, STC). Koncept STC-a čine sljedeće komponente: a) organizacija se povjerava jednoj zemlji članici koja STC potencijalno veže uz svoj nacionalni skup; b) Sastanak Vijeća EFMI-a je sastavni dio STC-a; c) temu konferencije izabire društvo organizator; d) u skladu s temom angažira se odgovarajuća EFMI-eva radna grupa; e) uglavnom sudjeluju pozvana predavanja; f) STC su male dvodnevne konferencije sa stotinjak sudionika.

Prva konferencija tipa STC održala se u Bukureštu/Rumunjska 2001. godine. Sljedeće su održane redom: Nicosia/Cipar 2002, Rim/Italija 2003, Muenchen/Njemačka 2004, Atena/Grčka 2005, Temišvar/Rumunjska 2006, Brijuni/Hrvatska 2007 and London/Velika Britanija 2008. 2009. godine STC organizira Antalya/Turska, 2010. Rejkjavik/Island a 2011. STC će se organizirati u Ptuj/Slovenija. Zbornike radova s konferencija obično publicira Springer u seriji 'Lecture Notes in Medical Informatics' i IOS Press u seriji 'Studies in Health Technologies and Informatics'. Izbor najboljih radova s kongresa MIE objavljuju se u posebnom broju časopisa International Journal of Medical Informatics i Methods of Information in Medicine, dok radove s STC-a objavljuje IOS Press kao kolekciju recenziranih članaka.

2. Pregled medicinskoinformatičkih časopisa u Europi

Medicinskoinformatička znanstvena i istraživačka produkcija u Europi raste i kvantitativno i u kvaliteti. Promovirajući visoko-kvalitetna istraživanja kao svoj glavni cilj EFMI ima dva visoko respektirana službena časopisa: International Journal of Medical Informatics (prijašnji naziv: International Journal of Bio-Medical Computing), s glavnim urednicima Charlesom Safranom i Janom Talmonom, te Methods of Information in Medicine, u kome je glavni urednik Reinhold Haux. EFMI objavljuje i više specijaliziranih časopisa koji pokrivaju cijeli spektar medicinskoinformatičkih potpodručja. Uobičajeni način diseminacije visoko kvalitetnih istraživanja odvija se putem zbornika radova prikazanih na MIE. EFMI ima i značajnu ulogu u edukaciji u harmonizaciji medicinskoinformatičkih potpodručja kroz organizaciju konferencija tipa STC. Oba zbornika radova, s MIE i STC, su recenzirana i objavljuje ih IOS Press (Amsterdam, Nizozemska). Uobičajeno, zbornici s MIE-a se indeksiraju u bibliografskoj bazi podataka Medline/PubMed. Mnogi medicinskoinformatički časopisi se objavljuju na lokalnim jezicima, a mnogi od njih uključuju sažetke na engleskom jeziku ili pak sadrže članke i na lokalnom i na engleskom jeziku. Sve u svemu, EFMI ima učinkovit način za diseminaciju medicinskoinformatičkih istraživanja koja se provode diljem Europe. Međutim, uz zbornike radova urednici časopisa imaju priliku da učine više u širenju znanja i boljem razumijevanju naše discipline.

3. Analiza izdavačkih aktivnosti unutar članica EFMI-a

Većinu nužnih informacije o EFMI-u može se naći na njegovim službenim mrežnim stranicama (www.efmi.org). Većina nacionalnih društava ima svoje mrežne stranice koje osiguravaju korisne informacije o publikacijama i sastancima u svojim zemljama. Naša se analiza zasniva na činjenicama objavljenim na EFMI-evim mrežnim stranicama pri čemu su te činjenice prikupljene od zemalja članica. Pretraživanjem Medlinea pomoću pojmova „EFMI“ i „European Federation for Medical Informatics“ dobiva se samo 21

članak 18 u medicinskoinformatičkim časopisima: Informatics in Primary Care (6), International Journal of Medical Informatics (3), Methods of Information in Medicine (2), Studies in Health Technologies (6) i 3 u ostalim časopisima. Autori članaka su: Blobel (2), De Lusignan (6), France (1), Horsch (1), McKeon (1), Masic (2), Michalas (2), Naszlady (1), Tallberg (1), Talmon (1), Trpisovský (1), van Goor (1), and Zvarova (1). Pretraživanje uz pojam „IMIA“ (radi usporedbe) daje 150 članaka. Pojam „AMIA“ rezultira s 244 članka.

Većina članaka nađenih s ključnom riječju „EFMI“ odnose se na aktivnosti radnih grupa PCI je najaktivnija (5,6) a relativno je malo inicijativa na razini federacije. Detaljnija analiza EFMI-evih mrežnih stranica pokazuje da samo 14 od 31 članice ima službene znanstvene i stručne medicinskoinformatičke časopise ili biltene (Tablica 1). Medicinskoinformatički časopisi publiciraju radove na engleskom jeziku a tri časopisa objavljuju radove na nacionalnim jezicima. Samo su tri časopisa koja izlaze isključivo u elektroničkom obliku. Učestalost publikacija registriranih s ISSN-om varira: jedan se publicira biennialno, dva kvartalno, jedan svaka dva mjeseca, pet časopisa izlazi 1-3 puta godišnje, dva mjesečno i jedan izlazi 9 puta godišnje. Glavna tema svih časopisa je medicinska informatika, no neki od njih se usredotočuju na informatička potpodručja. Većina medicinskoinformatičkih časopisa su međunarodne publikacije a tri su ograničena na nacionalnu razinu. Dva časopisa financira ministarstvo zdravstva i/ili druga javna ustanova dok se ostali financiraju preko pretplate.

4. Deklaracija urednika o svrsi i zadacima EFMI-evih časopisa

Glavni urednici medicinskoinformatičkih časopisa se pozivaju da zajednički kreiraju izjavu o svrsi i zadacima EFMI-evih časopisa. Vjerujemo da će zajednička izjava potaknuti širenje znanja i povećati razumijevanje medicinske informatike. Dodatno, namjera nam je proizvesti osnovni dokument s fundamentalnim principima s kojima se slažu svi urednici medicinskoinformatičkih časopisa. Treba definirati zajednički cilj i usuglasiti opseg i norme članaka koji će se objavljivati u

medicinskoinformatičkim časopisima. Razlozi za to su:

- Promocija uredničke izvrsnosti osnaživanjem i uporabom norma i smjernica (Smjernice Međunarodnog odbora urednika medicinskoinformatičkih časopisa; Odbor za etiku publiciranja; drugi relevantni dokumenti (7);
- Poboljšati kvalitetu znanstvenih studija objavljenih u EFMI-evim medicinskoinformatičkim časopisima postavljajući jednake norme za evaluacijske medicinskoinformatičke studije (službene EFMI-eve smjernice su već prihvaćene ali nema daljnjih koraka koji bi osigurali njihovu uporabu);
- Promovirati izvrsnost znanstvenog publiciranja jačanjem normi o obliku članaka (format rukopisa u obliku „camera-ready“; citiranje Vancouverskim stilom);
- Poboljšati širenje znanstvenih spoznaja u medicinskoinformatičkom području (prepoznavanje i širenje EFMI-evih istraživanja, edukacija, smjernice za kliničku praksu treba promovirati uzemljama EU);
- Ojačati suradnju urednika EFMI-evih medicinskoinformatičkih časopisa i dogovarati regularne sastanke grupe
- Istražiti potencijale razmjenjujući repozitorij recenzenata, informacije o procesu recenziranja: tehničke uredničke informacije, iskustva, inicijative, izdavačke resurse i alate (kao što je sustav otvorenih časopisa ili komercijalnih uredničkih rješenja);
- Upućivati članke koji ne odgovaraju određenom medicinskoinformatičkom časopisu da ih se objavi u drugima te održavati komunikaciju s urednicima tih časopisa;
- Zajednički se oglašavati kada je riječ o zajedničkim interesima i poboljšati suradnju između nacionalnih društava i EFMI-evih tijela;
- Promovirati europske inicijative stimulirajući publikacije i visokokvalitetna istraživanja.

Tab. 1. - EFMI member countries medical informatics journals.

Zemlja članica u EFMI	Časopis	Izdavač	Glavni urednik	Inicijalna godina	Frekvencija izlaženja (godišnje)	Način distribucije	Mrežne stranice	Jezik	Područje (MI i/ili drugo)	Recenzija	Financiranje
Austrija	Artificial Intelligence in Medicine	Elsevier	Klaus-Peter Adlassnig	1989 ISSN 1386-5056	9	email	da	engleski	MI i drugo	međunarodna	pretplata
Bosna i Hercegovina	Acta Informatica Medica	AVICENA, Sarajevo	Izet Masic	1993 ISSN 0353-8109	4	email	da	engleski	MI i drugo	međunarodna	pretplata
Hrvatska	Bilten HDMI	CSMI	Kristina Fišter	1992 ISSN 1330-0253	2	email	da	hrvatski	MI i drugo		Društvo za MI
	Medicinska informatika		Josipa Kern	1992 ISSN 1330-1799	biannually	email	da	engleski/hrvatski	MI	nacionalna	Ministarstvo znanosti, obrazovanje i sporta
Češka	Lékař a Technika	SBMIL	Karel Roubik	1970 ISSN 0301-5491	4	email	da	češki	MI i biomedicinsko inženjerstvo	nacionalna	javne institucije/pretplata
	European Journal for Biomedical Informatics	EuroMISE	Jana Zvarova	2005 ISSN 1801-5603	2	-	da	engleski+drugi	BMI	međunarodna	drugo
Danska	newsletter	DSMI	-	-	-	email	da	danski/engleski	MI	nacionalna	javne institucije
Francuska	Informatique et Sante	Springer-Verlag	Patrice Degoulet	-	2	email	da	francuski/engleski	MI	nacionalna	javne institucije/pretplata
Njemačka	Biometrie, Informatik und Epidemiologie in Medizin und Biologie	Urban & Fischer Verlag	-	-	-	-	da	njemački	MI i drugo	nacionalna	javne institucije/pretplata
Irska	Newsletter	-	-	-	2	-	da	-	-	-	-
Izrael	ILAMI journal and The Health Screen	-	-	-	2-3	email	-	-	MI i drugo	-	-
Slovenija	Informatica Medica Slovenica	SIMIA	Gaj Vidmar	1993 ISSN 1318-2129	6	-	da	engleski/slovenski	MI i statistika	međunarodna/nacionalna	Društvo za MI
Španjolska	I+S	SEIS	-	-	12	-	da	-	MI i drugo	-	-
Švedska	Computer Methods and Programs in	Elsevier	Torny Groth	1985 ISSN 0169-2607	2-3	papir/email/online	da	engleski	MI	međunarodna	pretplata
Ukrajina	Clinical Informatics and	UAMI	Oleg Yu Mayorov	2003 ISSN 1812-7231	4	email	da	ukrajinski/engleski	MI i drugo	međunarodna	javne institucije
Velika Britanija	Informatics in Primary Care	Radcliffe Publishing	Simon de Lusignan	1993 ISSN 1476-0320	12	papir + online	da	engleski	Informatika u PZZ i klinička informatika	međunarodna	PHCG BCS/pretplata
EFMI	International Journal of Medical Informatics	Elsevier	Charles Safran, Jan Talmon	1961 ISSN 1386-5056	6	papir/online	da	engleski	MI	međunarodna	pretplata
	Methods of Information in Medicine	Schattauer	Reinhold Haux	1962 ISSN 0026-1270	6	papir/online	da	engleski	MI	međunarodna	pretplata

Zahvala

Autori zahvaljuju na potpori Janu Talmonu sa Sveučilišta u Maastrichtu, Care and Public Health Research Institute, The Netherlands.

References

- [1] Masic I, Novo A.: History of medical informatics in Bosnia and Herzegovina. *Acta Informatica Medica*. 2007;15(1):49-61.
- [2] Collen M. F.: Fifty years in medical informatics. *IMIA Yearbook of Medical Informatics. Methods of Information in Medicine*. 2006;45(Suppl. 1):S174 -9.
- [3] Zvárová J.: Symposium on "Biomedical Informatics and Biomedical Statistics Education". *Methods of Information in Medicine*. 2006;45(3):281-2.
- [4] Zvárová J.: Biomedical Informatics Research and Education in the EuroMISE Center. *IMIA Yearbook of Medical Informatics. Methods of Information in Medicine*. 2006;45(Suppl. 1):S166-73.
- [5] de Lusignan S.: Introduction to the Primary Care Informatics Working Group of the European Federation for Medical Informatics (EFMI). *Informatics in Primary Care*. 2003;11(3):175-6.
- [6] de Lusignan S., Teasdale S.: Achieving benefit for patients in primary care informatics: The report of an international consensus workshop at MedInfo 2007. *Informatics in Primary Care*. 2007;15(4):255-61.
- [7] Miller R. A., Groth T/, Hasman A/, Safran C., Shortliffe E. H., Haux R., McGray A. T.: On exemplary scientific conduct regarding submission of manuscripts to biomedical informatics journals. *Methods of Information in Medicine*. 2006;45(1):1-3.

Contact

Izet Masic
 Medical Faculty
 University of Sarajevo
 Cekalusa 90/4
 71000 Sarajevo
 Bosnia and Herzegovina
 e-mail: imasic@lol.ba