

Automated Localization of Temporomandibular Joint Disc in MRI Images

Radim Burget, Petr Cika, Martin Zukal, Jan Masek

Abstract—This paper deals with localization of Temporomandibular Joint Disc (TJD) in Magnetic Resonance Images (MRI). Since the contrast of the TJD is quite low when compared to noise ratio when displayed using MRI, its detection is quite complicated. Therefore the method described in this paper are not not focused the disk itself but detect the most significant objects around TJD, which has usually much higher contrast. For the automatic TJD localization assessment, a training set containing 160 training samples (80 positive and 80 negative) were created and published and several approaches were examined to find the best method. The best results were achieved using support vector machine with Gaussian kernel, which achieved $98.16 \pm 2.81\%$ accuracy of detection. The creation of the training models for feature extraction and model evaluation was implemented with RapidMiner tool and the IMMI extension. The models created are published at the IMMI extension homepage and they can also serve as a guide to use of the IMMI extension.

Keywords—image processing, object detection, temporomandibular joint disc, MRI

I. INTRODUCTION

The Temporomandibular Joint Disc (TJD) is part of the joint of the jaw and is part of one of the most complex joint of the human body. This joint is composed of several parts, which interact with each other and realizes rotational translation of motion. Headache is a common symptom, but when severe, it may be serious problem. It is assumed that patients who visit dentists with headache should be often diagnosed with a Temporo-Mandibular Disorder (TMD), although many may have only migraine. Such a diagnosis is often caused by disc defect, which block the joint.

One of the most common methods for visualization in medicine is Sonograph, X-ray, Computer tomography (CT) and Magnetic Resonance Imaging (MRI) tomography. The reason why the MRI tomography was used in this paper is that MRI has certain advantages over the other methods - especially it does not use ionizing radiation.

Manuscript received May 18, 2011. This work was supported by the Grant MSM0021630513.

Radim Burget is with the Brno University of Technology, Department of Telecommunications Purkynova 118, 612 00 Brno, Czech Republic (phone: +420-541-149-176; fax: +420-541-149-192; e-mail: burgetrm@feec.vutbr.cz).

Petr Cika is with the Brno University of Technology, Department of Telecommunications Purkynova 118, 612 00 Brno, Czech Republic (phone: +420-541-149-205; fax: +420-541-149-192; e-mail: cika@feec.vutbr.cz).

Martin Zukal is with the Brno University of Technology, Department of Telecommunications Purkynova 118, 612 00 Brno, Czech Republic (phone: +420-541-149-205; fax: +420-541-149-192; e-mail: martin.zukal@phd.feec.vutbr.cz).

Jan Masek is student of the Brno University of Technology, Department of Telecommunications Purkynova 118, 612 00 Brno, Czech Republic (phone: +420-541-149-205; fax: +420-541-149-192; e-mail: xmasek10@stud.feec.vutbr.cz).

A 3D model of TJD can significantly help with diagnosis of the Temporomandibular Joint (TMJ). Unfortunately, creation of 3D models can be quite time consuming and also require knowledge of some 3D modelling tool. For this reason, this operation should be automatized with minimal required interaction with its user – i.e. most often a doctor. One of the approaches for an automatic 3D TJD reconstruction consists of 3 basic steps: 1) identification of TMJ presence and approximate locality, 2) segmentation of TMJ area and 3) construction of the 3D model. This paper focus on the first part of this chain, i.e. the localization of TJD in MRI tomography slices. The reason why steps 1) and 2) are both required is that the TJD when visualized using MRI has relatively low signal-to-noise ratio and surrounding object are often required to identify TJD.

The main contribution of this paper lies in examination of in total 19 different approaches for an automatic TJD detection in MRI images, their statistical evaluation and selection of the best method. For this purpose also a set containing 160 samples (80 images were without presence of TJD and 80 images with presence of TJD) was created. This training set and the evaluation processes created in Rapidminer were published¹. The best accuracy was achieved using a method based on Support Vector Machine (SVM) with the use of Gaussian kernel (accuracy: $98.16 \pm 2.81\%$).

The rest of the paper is structured as follows. The next section discusses a way how the training set was created. The section 3 describes the way of a feature extraction. The following section describes the way how the proposed models were statistically assessed and presents results achieved. The last section concludes the paper.

II. TRAINING DATA

As mentioned before, the problem of localization of TJD is that this object is hardly detectable in MRI images. This is caused by relatively low signal to noise ratio. Fortunately, there are several significant objects around TJD and therefore attention was focused on detection of the nearby objects. The procedure is very similar how a specialist proceeds to mark TJD in the images. On the basis of the TJD localization the steps 2) a 3) may follow as mentionde in the Introduction section.

The training set was created from a set of MRI images (individual slices) of 4 different persons. From the set of MRI slices 80 samples were selected, which contain TJD (so-called positive samples) and also 80 randomly chosen samples

¹<http://spl.utko.feec.vutbr.cz>

(so-called negative samples, which does not contain TJD). Since the slices from 4 patients would not contain enough information, several samples were duplicated and slightly modified (added noise, rotation, shift, scale, contrast, etc.). An example of images are depicted in the Fig. 1, where on the left half part of the figure positive samples are depicted and in the right half of the figure negative samples are depicted.

III. FEATURE EXTRACTION

In order to localize TJD in MRI images, several different features were extracted from the MRI images. All the features are so-called low level features. The aim of the extraction is to obtain numerical representation of each sample of dimension 150×150 pixels. On the basis of these features are decided, if the subfigure contain TJD or not. The features computed from MRI images are described below in the text.

The first method used was based on an extended version of Haar-like features introduced in [1], which was used with combination of AdaBoost algorithm described by the authors Viola and Jones [2]. The advantage of this approach is relatively high performance and good accuracy. Therefore it is suitable also for search in different scales. In case of TJD detection and the RMI images is expected, that scaling is not required since the scale of the TJD and its surrounding object is almost of a constant size.

Color and Edge Directivity Descriptor (CEDD) [3] is histogram based method, which incorporates color and texture information into histogram. Its advantage is relatively low computational power needed for their computation, especially in comparison with the needs of the most other MPEG-7 descriptors.

Auto color collerograms [4] approach is based on similarity measure idea. The similarity is measured between a reference (positive) image and the assessed image. The similarity between the reference sample and assessed sample is computed according to the probabilities of occurrence of the same color (or grayscale intensities) at a given distance from a given pixel. In comparison to histogram based methods this approach is more resistant to a color change.

General Color Layout is also a similarity measure approach. This approach attempts to overcome one of the drawbacks of histogram based approaches - i.e. that the location, shape and texture is lost. Images are partitioned into blocks and the average color of each block is stored, which correspond to some kind of a low resolution image. A drawback of the General Color Layout is that it is relatively sensitive to shifting, cropping, scaling, and rotation [5].

Color Layout Descriptor (CLD) is designed to capture the spatial distribution of color in an image. The method is composed of two main parts: grid based representative color selection and Discrete Cosine Transform (DCT).

Edge Histogram Descriptor (EHD) is a feature extraction method, which expresses only the local edge distribution in the image. Both CLD and EHD are descriptor described by the MPEG-7 standard in detail [6].

Fuzzy Color and Texture Histogram (FCTH) [7] in one histogram combines color and texture information. The advantage

of this methods is its resistance to deformations, adding noise and smoothing.

Fuzzy Color Histogram (FCH) is a histogram based approach. It considers the color similarity of each pixel's color associated to all the histogram bins through so-called fuzzy-set membership function [8].

Gabor similarity measure [9] is a similarity measure based function based on Gabor filters. The property of Gabor filters is that they are orientation dependent and this is an undesirable property for the purposes of images comparison. To cope with this problem the Gabor filters were modified in such a way that the modified function is free from the choice of angles.

Hue Saturation Value (HSV) color histogram [10] is another histogram based similarity measure function, which extracts features directly from the compressed and uncompressed (YUV) domain of an image. Its advantage is resistance to rotation, scaling, translation, and illumination correction.

JPEG coefficients histogram is also histogram similarity measure based approach. This approach utilizes DCT to measure similarity. The fact that many images are already stored in a form, where DCT coefficients are known (e.g. JPEG format) leads to a significant saving of computational power and it is also quite powerfull method for image comparison.

Joint Composite Descriptor (JCD) is a combination of CEDD and FCTH methods in texture areas and is described in detail in [11].

Tamura similarity measure is similarity measure approach based on texture features. It is based on research of a human perception of textures according to six basic textural features, namely: coarseness, contrast, directionality, line-likeness, regularity, and roughness. Its detailed description can be found in [12].

IV. EVALUATION

In order to evaluate the proposed model the cross-validation method was used. The Cross-validation is one of several approaches to estimate how well the learned model from a given training data is going to perform on as-yet-unseen data. This is performed by dividing data into two segments: one used to learn or train a model and the second to validate the model. In typical cross-validation, the training and validation sets must cross-over in successive rounds such that each data samples will be just once used as a validation data. Using this approach is especially suitable in case cases, when a training set is limited. The basic form of cross-validation is k -fold cross-validation, where commonly used value of k equals 10. The 10-fold cross-validation was used also in this paper.

A. Viola-Jones

The process of evaluation of the Viola-Jones AdaBoost algorithm [2] is depicted in the figure 2. The features used were the extended set of Haar-like features [1] counting 216 189 different Haar windows, where size of the window was 24×24 pixels. The process is performed as follows. First a set of training images is loaded. The training set is divided into two classes - the positive and negative samples as mentioned

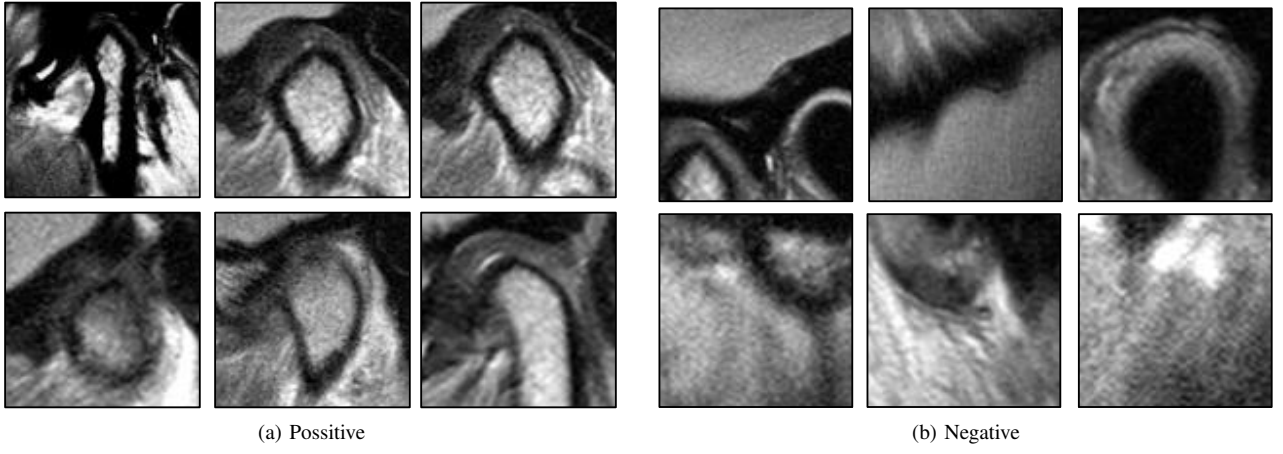


Fig. 1: Training example set.

before and as is depicted in the Fig. 1. Then integral images are computed. The integral images does not change information of the figure at all and its only purpose is just for faster computation of Haar features. The results are depicted in the Tab. I at row A and the achieved accuracy was $92.57 \pm 7.24\%$.

Since the results of the Viola-Jones algorithm is either yes with 100 % confidence or no with 100 % confidence, the Root Mean Squared (RMS) error was not expressed in the tab I.

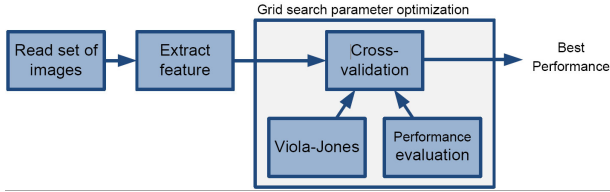


Fig. 2: Process of evaluation using Viola-Jones with extended set of Haar features. See Tab. I A.

B. Similarity measure functions

The others features are based on similarity measure techniques where the assessed image is compared to some reference image. The process of evaluation is depicted in the Fig. 3. Their advantage can be better resistance to adding noise, blur, rotation, scaling and other changes. To obtain the best results several learning algorithms were examined (Support Vector Machines, Decision Trees, and k -Nearest Neighbors) and with variety of different parameters for each learning algorithm. Searching for optimal parameters wasq performed using grid search method and only the best results were recorded. The results are depicted in the Tab. I from record B to record M.

As is from the tab I obvious, results obtained from similarity measure functions are significantly worse when compared to the results achieved with the Viola-Jones approach. The best accuracy were achieved using FCTH with accuracy ($82.65 \pm 10.29\%$). Unfortunately, the high value of the RMS error indicates that the accuracy estimation may not be too much reliable and size of the training set should be extended.

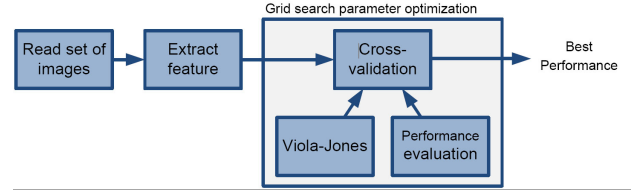


Fig. 3: Process of evaluation using image similarity measure methods. See Tab. I B-M.

C. Combination of several similarity measure functions

The last approach is depicted in the Fig. 3. This approach is based on combination of several similarity functions discussed in the text above and training a new training a new learning algorithm over the selected set of similarity functions. First the set of training set is loaded, then all the features with the exception of Haar-like features described in the section III are computed. The process of evaluation consist of combination of parameter optimization and feature selection. For the process optimization a grid search approach were used. For the feature selection the forward feature selection approach were used. The results were recorded into the table I in the rows N - U.

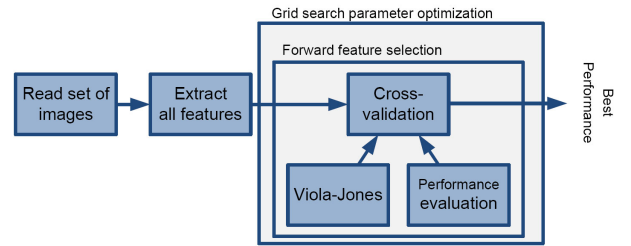


Fig. 4: Process of evaluation using image similarity measure methods. See Tab. I N-U.

What is obvious from the results achieved, accuracy of combined similarity measure functions is significantly higher when compared to use a single function. However, the detection accuracy is also higher when compared to Viola-Jones algorithm.

TABLE I: Comparison of different approaches.

ID	Algorithm	Accuracy	Negative class recall	Positive class recall	RMS error
A	Viola-Jones	92.57 \pm 7.24 %	95.00 %	90.12 %	N/A
B	Auto Color Correlogram measure	75.18 \pm 9.58 %	92.59 %	57.50 %	0.434 \pm 0.044
C	CEDD	75.07 \pm 7.52 %	82.72 %	67.50 %	0.428 \pm 0.042
D	Color Layout	80.81 \pm 7.36 %	77.78 %	83.75 %	0.390 \pm 0.036
E	Edge histogram	76.32 \pm 6.86 %	61.73 %	91.25 %	0.420 \pm 0.032
F	FCTH	82.65 \pm 10.29 %	75.31 %	90.00 %	0.368 \pm 0.061
G	Fyzyy Colour Histogram	72.54 \pm 10.61 %	71.60 %	73.75 %	0.430 \pm 0.051
H	Gabor	69.63 \pm 7.74 %	88.89 %	50.00 %	0.447 \pm 0.036
I	General Color Layout	72.13 \pm 8.52 %	74.07 %	70.00 %	0.417 \pm 0.039
J	JPEG Coefficient Histogram	77.06 \pm 6.68 %	81.48 %	72.50 %	0.411 \pm 0.035
K	Scalable Color	65.85 \pm 4.94 %	92.59 %	38.75 %	0.454 \pm 0.022
L	Simple Color Histogram	75.85 \pm 5.42 %	83.95 %	67.50 %	0.420 \pm 0.029
M	Tamura	68.86 \pm 9.42 %	83.95 %	53.75 %	0.458 \pm 0.041
N	SVM comb – radial kernel	97.50 \pm 3.06 %	97.53 %	97.50 %	0.182 \pm 0.058
O	SVM comb – dot kernel	97.50 \pm 3.06 %	97.53 %	97.50 %	0.182 \pm 0.058
P	SVM comb – linear kernel	96.25 \pm 4.15 %	95.06 %	97.50 %	0.169 \pm 0.070
Q	k-NN (k=1, Canberra measure)	95.04 \pm 6.12 %	96.30 %	93.75 %	0.153 \pm 0.162
R	Decision trees	93.20 \pm 4.29 %	92.59 %	93.75 %	0.234 \pm 0.106
S	SVM comb. – Gaussian kernel	98.16 \pm 2.81 %	97.53 %	98.75 %	0.305 \pm 0.038
T	SVM comb. – Neural network	95.07 \pm 4.56 %	96.30 %	93.75 %	0.230 \pm 0.099
U	SVM comb. – Anova	97.54 \pm 3.02 %	96.30 %	98.75 %	0.312 \pm 0.021

The highest accuracy was achieved using SVM with the Gaussian kernel. Unfortunately, the RMS is relatively high when compared to other learning algorithms. This can indicate overfitted learned model and thus also to a worse results on future un-seen data. From this point of view the SVM with the radial kernel and SVM with the dot kernel can seem to be also very interesting. The accuracy is slightly lower, however the RMS stay at lower values and therefore the confidence of the accuracy estimation is somewhat lower.

V. CONCLUSION

This paper deals with detection of TJD in order to facilitate 3D reconstruction of TJD on the basis of MRI images. Since the TJD has quite a low signal to noise ratio, the direct segmentation of the whole MRI image usually fail. Therefore the whole process is often divided into three steps: 1) TJD localization on the basis of the surrounding objects, 2) segmentation of the localized area and 3) 3D reconstruction. This paper was involved in the first part of this chain. For the purpose of the TJD localization, several approaches were examined: Viola-Jones algorithm, similarity functions and combination of several similarity function. The best results were achieved using SVM with Gaussian kernel but quite interesting results was achieved with SVM with radial kernel.

All the processes for feature extraction and for model evaluation (depicted in the figures 2, 3 and 4) were implemented using RapidMiner² tool and published at authors homepage of the image processing extension (IMMI)³.

²<http://rapid-i.com/>

³<http://spl.utko.feec.vutbr.cz/>

REFERENCES

- [1] Rainer Lienhart and Jochen Maydt. An extended set of haar-like features for rapid object detection. In *IEEE ICIP 2002*, pages 900–903, 2002.
- [2] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, volume 1, pages I-511–I-518, Los Alamitos, CA, USA, April 2001. IEEE Comput. Soc.
- [3] Savvas A. Chatzichristofis and Yiannis S. Boutalis. Ceddd: color and edge directivity descriptor: a compact descriptor for image indexing and retrieval. In *Proceedings of the 6th international conference on Computer vision systems, ICVS'08*, pages 312–322, Berlin, Heidelberg, 2008. Springer-Verlag.
- [4] Jing Huang, S. Ravi Kumar, Mandar Mitra, Wei-Jing Zhu, and Ramin Zabih. Image indexing using color correlograms. In *Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97)*, CVPR '97, pages 762–, Washington, DC, USA, 1997. IEEE Computer Society.
- [5] James Ze Wang, Jia Li, and Gio Wiederhold. Simplicity: Semantics-sensitive integrated matching for picture libraries. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23:947–963, 2001.
- [6] Philippe Salembier and John R. Smith. Mpeg-7 multimedia description schemes. *IEEE Trans. Circuits Syst. Video Techn.*, 11(6):748–759, 2001.
- [7] Savvas A. Chatzichristofis and Yiannis S. Boutalis. Fcthd: Fuzzy color and texture histogram - a low level feature for accurate image retrieval. In *Proceedings of the 2008 Ninth International Workshop on Image Analysis for Multimedia Interactive Services*, pages 191–196, Washington, DC, USA, 2008. IEEE Computer Society.
- [8] Guang-Hai Liu, Lei Zhang, Ying-Kun Hou, Zuo-Yong Li, and Jing-Yu Yang. Image retrieval based on fuzzy color histogram. *2008 International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, 43(7):75–78, 2008.
- [9] Challa S. Sastry, M. Ravindranath, Arun K. Pujari, and B. L. Deekshatulu. A modified gabor function for content based image retrieval. *Pattern Recogn. Lett.*, 28:293–300, January 2007.
- [10] Padmashri Suresh, R. M. D. Sundaram, and Aravindhan Arumugam. Feature extraction in compressed domain for content based image retrieval. *Advanced Computer Theory and Engineering, International Conference on*, 0:190–194, 2008.
- [11] Savvas A. Chatzichristofis, Yiannis S. Boutalis, and Mathias Lux. Selection of the proper compact composite descriptor for improving

content based image retrieval. In *The Sixth IASTED International Conference on Signal Processing, Pattern Recognition and Applications SPPRA 2009*, 2009.

- [12] Hideyuki Tamura, Shunji Mori, and Takashi Yamawaki. Texture features corresponding to visual perception. *IEEE Transactions on System, Man and Cybernetic*, 6, 1978.