International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:6, No:9, 2012

Face Image Coding using Face Prototyping

Jaroslav Polec, Lenka Krulikovská, Natália Helešová, Tomáš Hirner

***Abstract*—**In this paper we present a novel approach for face image coding. The proposed method makes a use of the features of video encoders like motion prediction. At first encoder selects appropriate prototype from the database and warps it according to features of encoding face. Warped prototype is placed as first I frame. Encoding face is placed as second frame as P frame type. Information about features positions, color change, selected prototype and data flow of P frame will be sent to decoder. The condition is both encoder and decoder own the same database of prototypes. We have run experiment with H.264 video encoder and obtained results were compared to results achieved by JPEG and JPEG2000. Obtained results show that our approach is able to achieve 3 times lower bitrate and two times higher PSNR in comparison with JPEG. According to comparison with JPEG2000 the bitrate was very similar, but subjective quality achieved by proposed method is better.

***Keywords*—**Triangulation; H.264; Model-based coding; Average face

I. Introduction

THE face is playing a major role in conveying identity and emotion. During our lifetime we can learn thousands of faces and identify similar faces at a glance even after years of separation. This skill is quite robust, despite large changes in the visual stimulus due to viewing conditions, aging, and distraction such as changes in hairstyle, glasses.

Face recognition has become an important issue in many applications such as security systems, credit card verification, criminal identification etc. Even the ability to merely detect faces, as opposed to recognizing them, can be important.

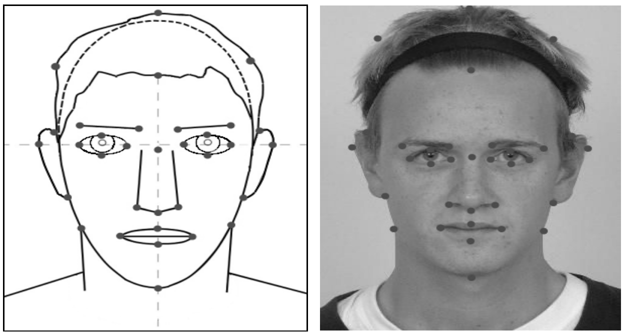
Although it is clear that people are good at face recognition, it is not at all obvious how faces are encoded or decoded by a human brain. Human face recognition has been studied for more than twenty years. Developing a computational model of face recognition is quite difficult, because faces are complex, multi-dimensional visual stimuli.

Therefore, face recognition is a very high level computer vision task, in which many early vision techniques can be involved. For face identification the starting step involves extraction of the relevant features from facial images. A big challenge is how to quantize facial features so that a computer is able to recognize a face, given a set of features. Investigations by numerous researchers over the past several years indicate that certain facial characteristics are used by human beings to identify faces [3].

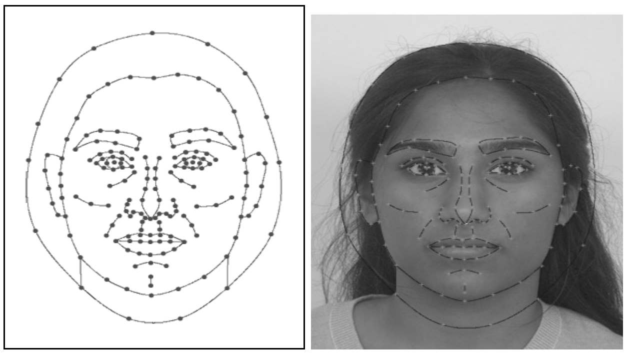
Jaroslav Polec is with the Institute of Telecommunication, Faculty of Electrical Engineering and Information Technology, Slovak University of Technology, Ilkovičova 3, 812 19 Bratislava, Slovakia (email: polec@ktl.elf.stuba.sk)

II. Prototyping

A prototype is an average of similar things [1]. In the 1800s, Francis Galton used multiple photographic exposures like the example on the Fig. 4 c) to make average faces. Using a sophisticated computer-graphic method called prototyping; we can make averages of groups of faces like the example on the Fig. 4 b).



a)



b)

Fig. 1 Example of a) simple proposed model [2]; b) sophisticated model [1]

First, we mark the same features on every face in the group. Different face models can be used. There are various numbers of points located corresponding to the characteristic features of human faces. There exist a lot of approaches to create ideal face model [4], [5].

Fig. 1 shows two models - simple with less number of features and more complicated with much higher amount of features.

The number of points (features), used for creating of average face, has no influence on encoding of face.

Therefore we can use for creating ideal average face in principle unlimited amount of points. Small amount of points, it means another face model, will be used for encoding process.

Then, we determine the average of every point to make an average position of each face feature (expressed by coordinates of points) and next, we warp every face into this average model. Or we make a single model, it means constant coordinates of features, and we warp every face into this face model (Fig. 2, Fig. 3).

a)

b)

Fig. 2 Group of faces from [1] deformed into a simple model [2]: a) originals, b) deformed

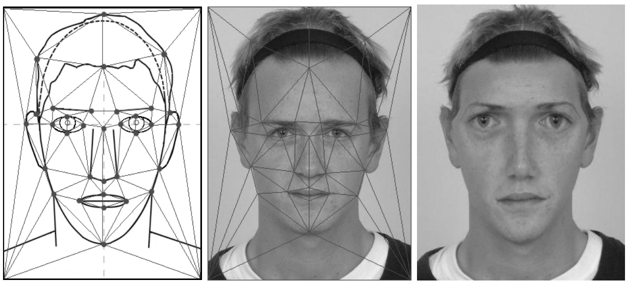


Fig. 3 Warping according to model [2]

Finally, we average the pixel values of all the warped faces to make the prototype face. We can obtain face in Fig. 4 a) by this way.

a) b) c)

Fig. 4 a) Average face obtained by simple model warping; b) average face obtained by warping into average shape [1]; c) average face obtained by simple averaging undistorted faces

III. Proposed Work

A. Preprocessing

We assume appropriate preprocessed prototype can well predict the encoding face. Preprocessed geometrically adapted (deformed, warped) face by a suitable prototype coded facial features. This means as the first step we will select prototype. The choice we make on the following criterions: age (baby, younger female, etc), ethnicity (East Asian female, white male, etc) and expression (angry female, sad male, etc), good categorization is in [1].

B. Encoder

If we use the obtained warped prototype face for predictive coding, we apply the following procedure:

1. Encoder sends features positions of encoding face and type of prototype to decoder. That is the reason why we use face model with small amount of features for encoding process. Decoder selects prototype from database and warps it according to features of encoding face.
2. Encoder selects chosen prototype from database and warps it according to features of encoding face. Encoder uses this warped face as the prediction of encoding face. The prediction is colored over quantized according to encoding face because of the luminance components and color variances between prototype and encoding face. Information about color shift is sent to decoder.
3. Encoder encodes encoding face with selected way by prediction. This means it is necessary to use encoders with frame prediction. This is the way how these two images become a video. The difference to video encoding is we don't need to encode and transmit first frame (prediction). The condition is both encoder and decoder have to have the same database of prototypes, what is very similar to higher level vector quantization.

a) b)

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:6, No:9, 2012

c) d)

Fig. 5 a) Original (coded) face, (b, c, d) example of prototypes and models of different groups according to warped coded facial features (a) obtained by the model Fig. 1 (a) (b) from a prototype "african male" [1], (c) of the prototype made of images of [1] african males and females, (d) neprototyped face from images of faces [1] different ethnicities [2]. Prototypes for production (b) and (c) were obtained by the model Fig. 1 (b), neprototypová face for production (d) was obtained by the model Fig. 1 (a)

It would be possible to use the direct difference between prediction and encoding picture and to encode the obtained result by JPEG [6] or JPEG2000 [7]. The disadvantage of this procedure is that the standard image with bit rate of 8 bpp (or 24 bpp for RGB images) obtained bit rate 9 bpp (27 bpp for RGB images). Currently available coders JPEG and JPEG2000 do not work with such a bit rate, it would be necessary to use specially prepared coders. This limits the usefulness of mentioned procedure.

Prediction is used in video coders as a basic precondition for achieving a high compression [8]. Current video encoders have in comparison with encoders of static image the advantage they are able to improve imperfect prediction by motion vectors. Moreover the motion compensated blocks from prototype can be directly copied in decoder without encoding their content. Based on previous consideration we propose to use video encoder as follows:

1. Warped prototype will be placed as first frame. The type of frame will be I frame with the best possible quantization parameter.
2. Encoding image will be second frame. It will be P frame with chosen quantization parameter.
3. Information about features positions, color change and data flow of P frame will be sent to decoder.

IV. Experiment And Results

For the experiment we used the H.264 Video Encoder (in Fig. 9 h.264, P) [9], picture of a man of African ethnicity [1], which we chose as a prototype "African male" [1]. The results were compared with results achieved by coders JPEG and JPEG2000 [9] in standard mode - the face is encoded as single static image and by H.264 encoder, which wasn't use in standard way, where the face was coded as video with one frame (in Fig. 1 h.264, I).

Fig. 6 shows an original image, which should be encoded and warped prototype according the features of encoding face. Fig. 7 and Fig. 8 display the obtained decoded image for encoders of static images - JPEG and JPEG2000 and for H.264 video encoder obtained by encoding video with one single frame (frame type I) and by proposed approach.

From the view of subjective evaluation the best result was achieved by proposed method. From the view of bitrate comparison, JPEG needed 0,14 bpp, the rest of simulated methods only 0,05 bpp. According to PSNR, JPEG reached 16,8 dB, JPEG 31,7 dB, H.264 coded as I image 33,3 dB and , H.264 coded as P image, where warped prototype was used as a prediction 34,6 dB.

a) b)

Fig. 6 a) Original image; b) average face warped by coded image using model from [1]

a) b)

Fig. 7 a) decoded image, JPEG, 0,14 bpp, PSNR 16,8 dB.; b) decoded image JPEG2000, 0,05 bpp, PSNR 31,7 dB

a) b)

Fig. 8 a) decoded image, H.264 coded as I image, 0,05 bpp, PSNR 33,3 dB; b) decoded image, H.264 coded as P image, where warped prototype was used as a prediction, 0,05 bpp, PSNR 34,6 dB

Comparison of PSNR results of the decoded image depending on the bitrate listed in Fig. 9.

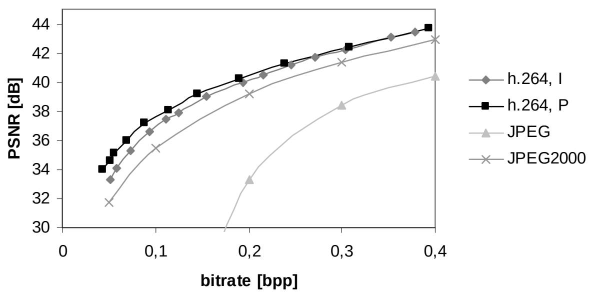


Fig. 9 PSNR performance of H.264, JPEG and JPEG2000 coder

V. Conclusion

In this paper, we have presented a new procedure for image coding using face prototyping. Proposed method makes a use of features of video encoders, like motion estimation, with are not included in encoder of static images.

The principle of proposed approach is very simple - at first encoder selects appropriate prototype from database and warps it according to features of face, which should be encoded. Subsequently the warped prototype is set as first frame (frame type I) and image of encoding face as second frame (frame type P). Information about features positions, color change, selected prototype and data flow of P frame will be sent to decoder. Decoder selects prototype from database and warps it according to features of encoding face.

We verified the effectiveness of proposed method with H.264 video encoder by test experiment and the obtained results were compared with encoders of static images - JPEG and JPEG2000 and with the result achieved by H.264 video encoder, where the face was coded as video with single frame.

Obtained result show that proposed methods give the best result according to bitrate, PSNR and subjective quality. Visually decoded image by our proposed methods is most similar to the original one.

As shown in Fig. 9 H.264 video encoder gives the best result, better one for proposed method with prediction. JPEG and JPEG2000 didn’t reach the results obtained by our method, JPEG is significantly worse in terms of bitrate and PSNR.

Acknowledgment

Research described in the paper was financially supported by the Slovak Research Grant Agency (VEGA) under grants No. 1/0961/11 and 1/0602/11.

References

[1] L. Debruine, B. Jones, FaceResearch.org. University of Aberdeen in Scotland. 2010. http://www.faceresearch.org/demos/average

[2] P. Weiss, *Coding of Facial Images by Triangulation*, Thesis, SUT Bratislava, 2010

[3] M. Turk, A. Pentland, "Eigenfaces for Recognition," *Journal of Cognitive Neuroscience*, March 1991

[4] [12] E. Šikudová, M. Gavrielides and I.Pitas, “Extracting semantic information from art images,” *Computer Vision and Graphics : International Conference, ICCVG 2004*, pages 394-399, Warsaw, 2004.

[5] E. Šikudová, “Comparison of color spaces for face detection in digitized paintings,” *Spring Conference on Computer Graphics SCCG 2007*, pages 135-140, Bratislava, 2007.

[6] Joint Photographic Experts Group, oficial webpage, http://www.jpeg.org/jpeg

[7] Joint Photographic Experts Group, oficial webpage, http://www.jpeg.org/jpeg2000

[8] J. Polec, J. et al., *Selected Methods for Data Compression*, Bratislava: UK, 2000. 196 p.

[9] *Image and Video Compression Learning Tool VcDemo*, Signal & Information Processing Lab, Delft University of Technology, http://siplab.tudelft.nl/content/image-and-video-compression-learning-tool-vcdemo