Robert KRUPIŃSKI, Przemysław MAZUREK, Konrad DERDA

West-Pomeranian University of Technology Szczecin

doi:10.15199/48.2015.02.11

Small-Size Skin Features for Motion Tracking

***Abstract**. Motion tracking systems for face or hands are usually based on tracking markers. Using markers for faces is inconvenient and there is small number of natural existing face features for precised movement tracking. Using high-frame rate and high resolution camera or cameras skin surface tracking is possible. The requirements for such system based on reduced number of markers and chromatic skin features for light sensitivity reduction are investigated in this paper.*

***Streszczenie.** Systemy śledzenia ruchu twarzy lub rąk zwykle bazują na śledzeniu znaczników. Wykorzystanie znaczników dla twarzy jest niewygodne, jednak możliwe jest wykorzystanie istniejących cech twarzy. Wykorzystanie szybkiej i wysokorozdzielczej kamery lub kamer umożliwia śledzenie cech skóry. Wymagania dla systemów bazujące na zredukowanej liczbie znaczników oraz cechach skóry są rozpatrywane w artykule. (**Mało-rozmiarowe cechy skóry dla śledzenia ruchu**)*

Keywords: Pattern Recognition, Tracking, Virtual Cinematography, Motion Capture.
Słowa kluczowe: Rozpoznawanie obrazów, Śledzenie, Kinematografia wirtualna, Przechwytywanie ruchu.

Introduction

Facial or hand positions are usually estimated using features like eye, lips corners and body part silhouettes. Markers placed on skin or projected light patterns are used for more precised estimation of 3D data. Using skin features is also interesting because skin has a lot of different scale features. The tracking such features requires fast and reliable algorithms but two aspects are especially important. The first problem is high image resolution because skin features should be acquired by camera but they are small (millimeter order). The second problem is related to the method of estimation, because such features are not points like and the estimation of position should be based on pixels area. Markers and existing face or hands features are very reliable and a lot of algorithms and systems were proposed [1,2,3]. Most of them are light condition independent and used as a man-machine interfaces and in biometric identification applications.

There is also another group of applications where very precise measurements are necessary. For virtual cinematography usually marker based tracking systems are used due to high reliability [4]. Motion capture systems are used with sets of synchronized cameras for face features tracking. Motion capture systems could be used for full body, but for additional face tracking system is used also. Manually controlled cameras (long focal length) are applied for the face image of principal actors acquiring. Hundreds of markers are used for face tracking in both cases. Placing markers on actors's face is time consuming because markers should be located exactly in the same places every time (every motion capture session). Alternative approach uses depth camera [5].

The approach based on hybrid motion capture method that is based on markers or existing large scale features and additionally uses small scale skin features is proposed in this paper. Large scale features are used for the reduction of computation cost for raw estimation of skin movements. Small scale skin features are used for the reduction of amount of markers. Such approach should reduce cost of motion capture maintenance with preserving tracking quality. Such approach improves safety, especially for eyes because markers (small balls) can injure eyes additionally. Overall hybrid process based on markers or existing face/hand features (marked as dots) and more precise skin movements is shown in Fig.1.

Color Spaces for Motion Estimation

Detection of skin area using color camera and chroma-based color space like HSV or Lab is possible. Placing actor in the front of contrast background, like black wall and floor, is typical technique for the estimation of actor's region in motion capture systems and could be used for the reduction of computations.

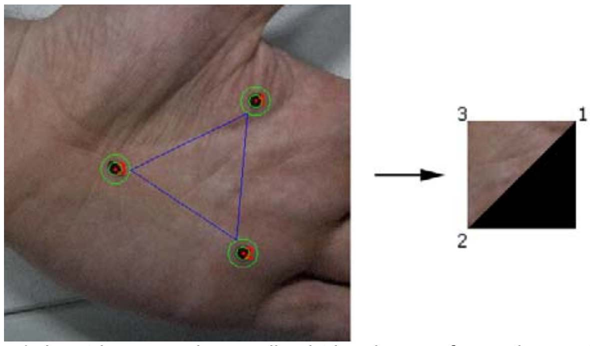

Fig.1. Input image and normalized triangle area for motion vectors estimation

Skin area can be estimated using saturation and hue values because they are common for all people [6,7]. This technique is very often applied, because is light independent (but sensitive for light color changes and color temperature).

Human skin is no smooth but bumpy due to e.g. wrinkles, fingerprints so luminance based tracking is possible. Such features are well visible for low-angle lighting due to light introduced shadows and contrast. The luminance based method is not reliable if variable lighting condition occurs. Low-angle lighting is not available every time and cannot be used for round objects (including head, hand, and fingers) - only part of them will be illuminated by such desired light. The luminance method is interesting for 3D scanning for fixed position of e.g. head using controlled light environment. Multiple approaches for the estimation of skin movements depending on color space are possible. The simplest, because it is direct method, is the utilization of one of the subspaces (channels) selected from RGB space. Every channel of RGB consists of partial information about saturation, hue and luminance that is the main drawback. RGB channels are sensitive to variable lighting and they are not useful like other color spaces, but could be used as a reference.

The second approach is based on Lab space [8]. Luminance channel has a lot of small scale features, but it is light condition dependent. Two last channels ('a' and 'b') are artificial channels without intentional naming or physical correspondence each, but both are chroma channels.

Because they are related to chromatic features of skin, they could be used for the estimation of skin movements.

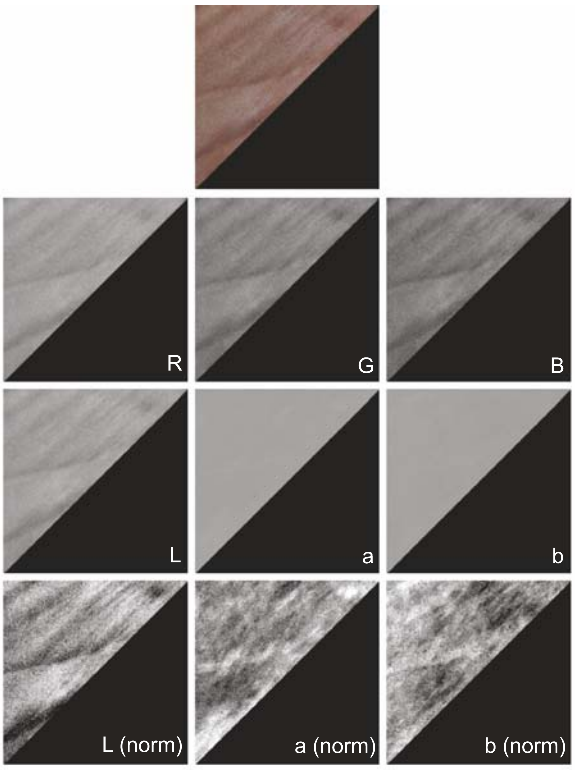

Fig.2. Original color image and set of color channels

The difference between visible features for RGB and Lab spaces are visible in Fig.2. RGB channels are similar to the luminance channel and 'ab' channels are very flat (they have low contrast). This is typical behaviour for skin samples and both flat channels are results of low variance for the chroma and spatial similarity for skin. After normalization, for contrast enhancement, the chroma features are well visible ('norm' row in Fig.2). Using 24-bit RGB images 'ab' channels are based on a few bits so high quality camera should be used.

Channel-Based Motion Estimation

Normalization of input measurements is used before motion estimation. Three neighbourhood points represent corners of triangle, where the markers are placed. Both triangles (two from consecutive frames) are normalized so they are equal size. They are affine transformed and both are perpendicular additionally. Such approach needs additional computations, but simplifies further computation efficient implementations based on SIMD (Single-Instruction Multiple-Data) processor architectures and GPGPUs (General Purpose Graphical Processing Units).

The window approach is applied, for the motion estimation, where the first window is fixed for first image and the second is moving for second image. As a metric the mean absolute error (MAE) is used and minimal value.

(1) $$E_{i,j} = \frac{1}{WH}\sum_{m}\sum_{n}\left|W_1(x_1+m, y_1+n) - W_2(i+m, j+n)\right|$$

(2) $$E_{\min} = \min_{i,j}\left(E_{i,j}\right)$$

where: $W_1(x_1,y_1)$ - image window for first image located at position x_1, y_1, $W_2(i,j)$ - image window for second image located at position i, j; W - window width, H - window height, N - number of windows.

The synthetic perturbed images based on real skin samples are used for performance test (dependent on channel). Such approach allows the comparing of estimated and assumed motion vectors.

The results for RGB and Lab channels are shown as example in Fig.3. It is assumed 20x20 mask, pixel step movements resolution and the maximal movements is 10 pixels in all directions. This block size gives good and comparable results for all cases, but for particular mask size. Median filters are used for filtering motion vectors and the 15x15 filter mask size is assumed. The results depend on the assumed window size. Large window sizes give poor fitting, because local skin movements (deformations) are observed. Very small windows give local assignments to the local features only and fitting fails also.

The estimation of optimal window size and its sensitivity is crucial for the motion estimation systems. High sensitivity will disqualify considered approach (for selected channel) so they are crucial tests.

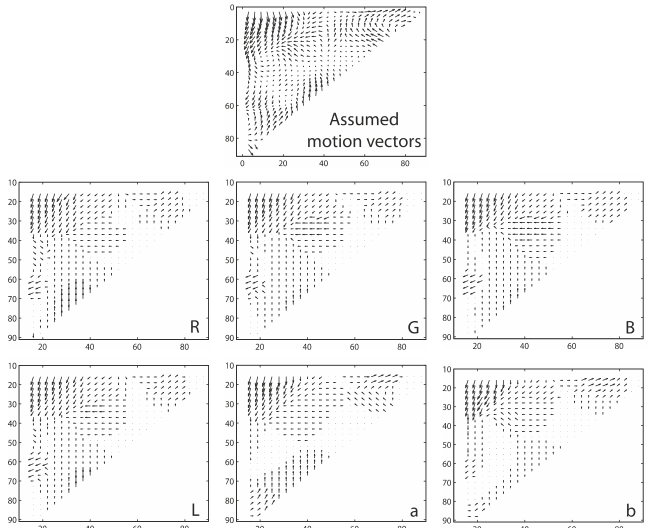

Fig.3 Original motion vectors, estimated RGB and Lab channels movements example for optimal size

The results for RGB channels are shown in Fig.4. Blocks sizes have minimal sensitivity for 8-25 range.

The results for Lab channels are shown in Fig.5. Blocks sizes have minimal sensitivity for 8-35 range for 'ab' channels. There is a larger increase in the MAE error for the luminance channel together with increasement masks size because of the inclusion of larger areas for comparison and where new values do not add more information but rather uncertainty. Obtained optimal values (ranges) could be applied for the motion estimation of the skin features.

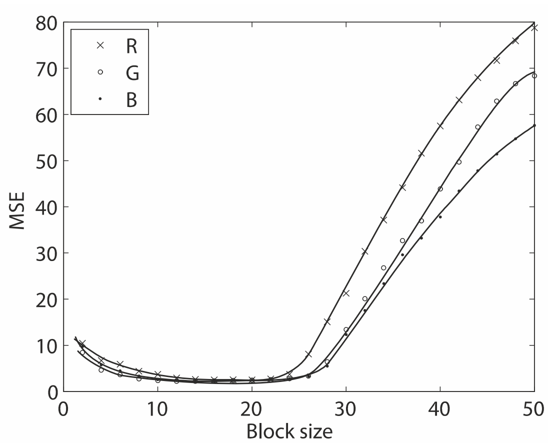

Fig.4 MAE for RGB blocks depending on size

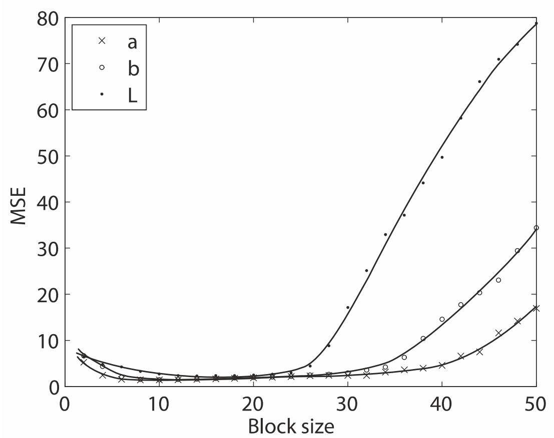

Fig.5 MAE for Lab blocks depending on size

Motion Estimation for Large Mask Size

In next example 34x34 mask size is assumed - largest possible. According to Fig.4 and Fig.5 such size is optimal for 'ab' channels but not for RGB and luminance. The results are shown in Fig.6 and the best results are for 'ab' channels.

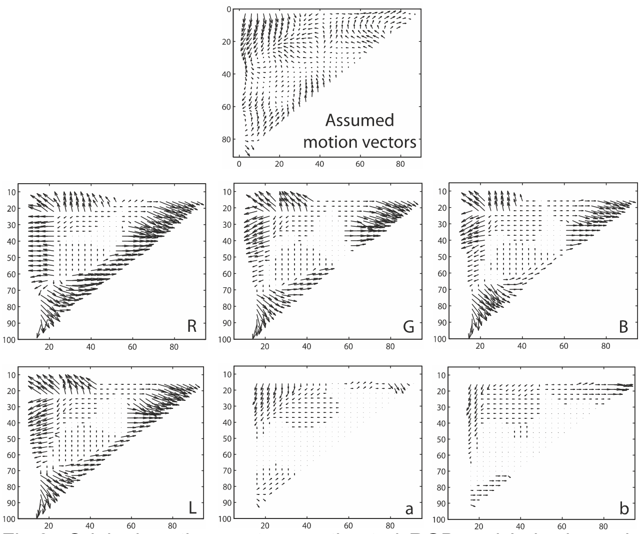

Fig.6. Original motion vectors, estimated RGB and Lab channels movements example for non-optimal size

Particular area has different size depending on skin and camera orientations. Less sensitive channels ('a' and 'b') should give better results because such relation is not known or known with some certainly.

The boundary vectors are constrained and cannot point outside the considered texture. The windows at the edges for the comparison are reduced to the valid size covering points from the normalized image. The values outside are not taken into account.

Conclusions

Proposed solution is based on the hybrid tracking, that allows the application of reduces set of the markers - painted dots on skin or small balls attached to the skin surface. The area between markers is tracked using small features of skin so dense motion flow could be obtained, desired by e.g. face tracking system.

In the article, the application of Lab color space, especially, 'a' and 'b' channels for the improvement of estimation human skin movements and deformations was presented.

These channels were proved to be most appropriate for the tracking. This type of movements estimation was analyzed under the varying the block size. The 'a' and 'b' method is less sensitive to the block size which is important when the skin has different orientation in 3D space.

The implementation of motion capture system based on the tracking small skin features requires the estimation of the tracked area. The area of analysis (e.g. human head) could be estimated using background modelling [9,10]. Multiple marker tracking requires dedicated algorithms and acquisition systems [11,12]. Texture analysis techniques for human tracking, related to the clothing are recently considered in [13,14].

This work is supported by the UE EFRR ZPORR project Z/2.32/I/1.3.1/267/05 "Szczecin University of Technology - Research and Education Center of Modern Multimedia Technologies" (Poland).

REFERENCES

[1] Wechsler H., Reliable Face Recognition Methods. System Design, Implementation and Evaluation. Springer (2006)
[2] Zhao W., Chellappa R. (eds.): Face Processing: Advanced Modeling and Methods. Academic Press (2005)
[3] Li S.Z., Jain A.K. (eds.): Handbook of Face Recognition. Springer (2005)
[4] Moeslund T.B., Granum E.: A Survey of Computer Vision-Based Human Motion Capture. Computer Vision and Image Understanding, Vol. 81, 231-268 (2001)
[5] Hachaj T., Ogiela M.R., Piekarczyk M., Real-time Recognition of Selected Karate Techniques using GDL Approach, Advances in Intelligent Systems and Computing, Vol. 233, 99-106 (2014)
[6] Vezhnevets V., Sazonov V, Andreeva A.: A Survey on Pixel-Based Skin Color Detection Techniques, Proc. Graphicon-2003, 85-92, Moscow, Russia, September (2003)
[7] Jones M.J., Rehg J.M.: Statistical Color Models with Application to Skin Detection, International Journal of Computer Vision, Vol. 46 No. 1, January 2002. ACM (2002)
[8] Fairchild M.D.: Color Appearance Models. Willey (1997)
[9] Frejlichowski D., Gościewska K., Forczmański P., Nowosielski A., Hofman R.: The removal of false detections from foreground regions extracted using adaptive background modelling for a visual surveillance system. LNCS Vol. 8103, 253-264 (2013)
[10] Frejlichowski D., Grzegorzewicz K.: An approach to automatic detection and extraction of regions of interest in still images. AICS Vol. 184, 3-10 (2013)
[11] Mazurek P.: Estimation of state-space spatial component for cuboid track-before-detect motion capture systems, LNCS Vol. 5337, 451-460 (2009)
[12] Mazurek P.: Estimation track-before-detect motion capture systems state space spatial component, LNCS Vol. 4673, 149-156 (2007)
[13] Forczmański P., Czapiewski P., Frejlichowski D., Okarma K., Hofman R., Comparing clothing Styles by Means of Computer Visio Methods, LNCS Vol. 8671, 203-211 (2014)
[14] Okarma K., Frejlichowski D., Czapiewski P., Forczmański P., Hofman R., Similarity Estimation of Textile Materials Based on Image Quality Assessment Methods, LNCS Vol. 8671, 478-485 (2014)

***Authors**: dr inż. Przemysław Mazurek, Zachodniopomorski Uniwersytet Technologiczny w Szczecinie, Katedra Przetwarzania Sygnałów i Inżynierii Multimedialnej, ul. 26. Kwietnia 10, 71126 Szczecin, E-mail: robert.krupinski@zut.edu.pl, dr inż. Przemysław Mazurek, Zachodniopomorski Uniwersytet Technologiczny w Szczecinie, Katedra Przetwarzania Sygnałów i Inżynierii Multimedialnej, ul. 26. Kwietnia 10, 71126 Szczecin, E-mail: przemyslaw.mazurek@zut.edu.pl, Konrad Derda, Zachodniopomorski Uniwersytet Technologiczny w Szczecinie*