Evaluation and comparison of tracking methods using meeting omnidirectional images

Igor Potucek, Vitezslav Beran, Stanislav Sumec and Pavel Zemcik

Department of Computer Graphics and Multimedia, Faculty of Information Technology,
Bozetechova 2, Brno 62500, Czech Republic
{potucek, beranv, sumec, zemcik}@fit.vutbr.cz

Visual cues, such as gesturing, looking at each other or monitoring each others facial expressions, play an important role in meetings. Such information can be used for indexing of multimedia meeting recordings. These situations are strongly focused nowadays. The omnidirectional system usage in such situations brings many advantages as portability, easy installation, large field of view, low cost etc. That is why we choose such scenarios for testing the omnidirectional system. Information about differences between omni-directional and classical images both for human presentation and tracking purposes is needed. We try to compare two different tracking methods on the various video sequences. The results of the tracking methods can help to demonstrate the benefits or drawbacks of the omni-directional system. The evaluation scheme was developed to bring us the aspects which affect the vision algorithms for detection and tracking of human bodies.

Introduction

Monitoring of meetings usually requires several cameras to capture the whole scene with each participant. Conventional cameras have a relatively narrow field of view. It could for instance use a pan-tilt-zoom mechanism to aim the camera in different directions, or it could rotate its body. Recently, an increased interest in omnidirectional vision for applications not only in robotics could be noted. Technically, omnidirectional vision, sometimes also called panoramic vision, can be achieved in various ways. Examples are cameras with extremely wide angle lenses ("fish-eye"), cameras with hyperbolically curved mirrors mounted in front of a standard lens (catadioptric imaging), sets of cameras mounted in a ring- or sphere-like fashion, or an ordinary camera that rotates around an axis and takes a sequence of images that cover a field of view of 360 degrees. Omnidirectional vision provides a very large field of view, which has some useful properties. For instance, it can facilitate the tracking of moving objects in the scene.

Despite of the wide usage, we will aim to monitor participant activity at live meetings, which are important part of everyday human social life as was stated above. It is convenient to retain information generated in the meetings for later use. The traditional approach of manual transcription is time consuming, modern technology can help to automate meeting recording and processing. The goal is to monitor participant activity in the whole scene and extract relevant information about movement, gestures, pointing, voting and other human actions.

The comparison of tracking methods to different types of source data is presented in this paper. These tests are performed on the data set containing recordings from standard perspective cameras and omnidirectional system. The evaluation protocol and person tracking methods are suggested.

Catadioptric system description

In this work, the system consisting of an ordinary perspective video camera equipped with a hyperbolic mirror is studied. Such system allows capturing of a large portion of the space angle, usually 360x105 degrees field of view.

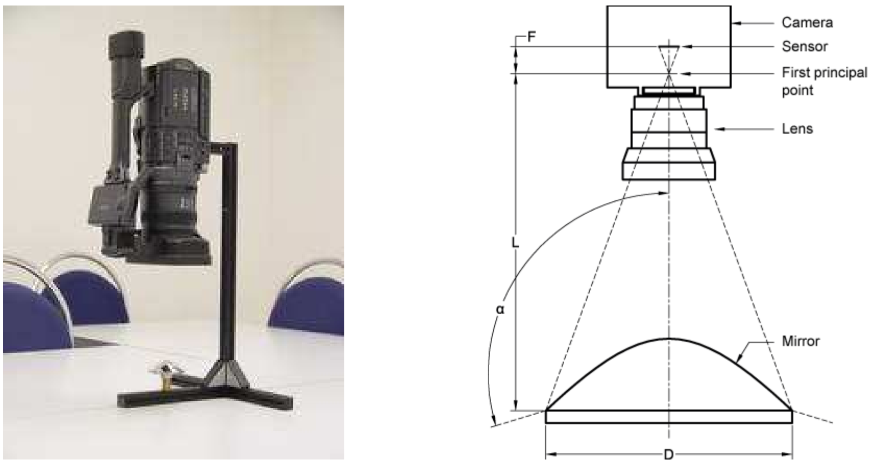

Fig. 1 a) Demonstration set-up with conventional camera and mirror holder b) optimal camera and lens configuration

Two setups exist for capturing such images: mirror above/below the camera. The mirror above the camera is usually used for mobile robots for terrain observing. The second setup is more suitable for capturing people sitting around the table. Before further processing, each image is transformed to a standard perspective or panoramic view[2][3]. The omni-directional image has different features than image captured by standard camera. Vertical resolution of the transformed image has usually non-uniform distribution. The circle which covers the highest number of pixels is projected from the border of the mirror, which means that the transformed image resolution is decreasing towards the mirror center. If the image is to be presented to a human, it needs to be a perspective/panoramic image so not to appear distorted. Other issues become more important, when we want to process the image further, such as spatial resolution, sensor size and ease of mapping between the omni directional images and the scene.

Detection and tracking methods

The aim of the system is intended to incorporate various methods into single tracking system of human body parts and to evaluate different method of combinations. Two approaches were tested to detect heads/faces and track them. The first approach is based on the pure skin color segmentation; face detection and tracking using the movement prediction. The second approach is using of the head detection based on the skin color and background subtraction. KLT tracker then provides the tracking.

Skin color detection was used as the basic method for image segmentation. The appearance of the skin-tone color depends on the lighting conditions. Hence normalized rg-color space is used, which is good solution for the problem of varying brightness. Normalized rg-color are computed from RGB values. The Gaussian color model was trained to achieve better results which are needed for image segmentation. Skin colored blobs are obtained by connected component analysis and morphological operations. The face detection is then applied only on the detected skin colored areas for increasing of the speed of the whole algorithm. The detected skin colored objects, which are recognized as faces, are then tracked by using movement prediction. The information used for this purpose is only the movement. The motion equations based on basic physics are used for estimating of a new position from previous object movement. Prediction of object position is therefore based only on the positions in past frames. The face detection algorithm is then applied only on the detected skin colored areas, which significantly increases the speed of the whole algorithm.

The face detection algorithm is based on the well known AdaBoost learning algorithm. The face classifier is constructed as a linear combination of several weak classifiers (e.g. a simple perceptron) built on features issued from the AdaBoost algorithm. In our case, the simple rectangle image/facial features are replaced by more complex Gabor wavelets and a modified confidence-rated AdaBoost algorithm is used for learning. Each weak classifier is composed of the Gabor wavelet and a decision tree whose output determines "confidence" that the input image is a face. The training algorithm considers fact that there is much more non-face regions then face regions in an image. Therefore, during the classification, the non-face regions are recognized and rejected faster. Face detector is trained on normalized face images (24x24 pixels). Firs, the input image is sub-sampled and rotated. Then, the face detection is performed, scanning sub-sampled images by the normalized window. After processing the input image, all possible occurrences of faces are grouped by the help of a clustering algorithm. This helps to stabilize position of detected faces over a video sequence and it also imporoves overall detection rate. Such algorithm is able to detect faces rotated along the eye-view axis and partially rotated along the vertical/horizontal axis.

The second tracking method is based on the public domain KLT feature tracker, which uses an image pyramid in combination with Newton-Raphson style minimization to efficiently find a most likely position of features in a new image. We embedded both flocking behavior and color cue into the tracking system. We used RG color model as color cue that could be either predefined or trained when tracker is placed on an object. We use this model in order to discard all features whose color does not match expected object color. This color cue in combination with the flock compactness criterion almost eliminates feature drift to background and nonstationary objects in the scene. Tracker is also resistant to partial occlusions. For the head detection, we use the presumption that faces correspond to compact ellipse-like shapes with distinctive axis aspect ratio in the mask. The skin color analysis, background subtraction and connected component analysis are used to extract suitable object for head detection.

The method based on progressive background model improvement was used. Model improvement is done by accumulation of RGB pixel values of current frame in model buffer. Only those pixels evaluated as background are updated. The spatial component analysis by statistical moment calculation is used to distinguish between the heads and other skin colored human parts. Note that results are used for tracker initialization, not tracking itself, though it may be suitable even to perform the tracking itself. The first method is based purely on the skin color detection and therefore is gives us good results only under certain lighting conditions. The negative aspect of our face detector is strong computation dependency on the Gabor wavelet feature evaluation and therefore it cannot be used in real-time applications. The speed of this algorithm on Athlon 64 3500+ computer is about 0.2 frames per second.

The tracker containing KLT method is not so sensitive to lighting conditions because it uses the background subtraction for the image segmentation in addition. On the other hand, the FP(false positive) rate is quite high – around 30 % because hands are often misinterpreted as heads. The FP rate could be reduced using additional topological knowledge about the scene and temporal correspondence, or by using a face detector. The KLT tracking algorithm with the head detection achieves approximately 17 frames per second on Athlon 64 3500+ processor, when it tracks two persons in standard DV sequence.

Evaluation procedure

To objectively compare the tracking and detection methods on several video sources, we must first define a common evaluation. The task of evaluating tracker performance[4] was broken into evaluating two tasks: predicting of the correct number and placement of objects in the scene (referred to as configuration), and checking the consistency with which each tracking result (or estimate, E) assigns identities to a ground truth object (GT) over its lifetime (referred to as identification). Several metrics are defined below to evaluate these tasks. Each of these measures depends on information derived from the fundamental coverage test. The coverage test determines if a GT is being tracked by an E; e.g. if the E is tracking the GT, it reports the quality of the tracking result. For a given tracking estimate E_i and ground truth GT_j , the coverage test measures the overlap between the two areas using the F-measure $F_{i,j}$

$$F_{i,j} = \frac{2\alpha_{i,j}\beta_{i,j}}{\alpha_{i,j} + \beta_{i,j}} \qquad \alpha_{ij} = \frac{|E_i \cap GT_j|}{|GT_j|} \qquad \beta_{ij} = \frac{|E_i \cap GT_j|}{|E_i|} \qquad (1)$$

where recall (α) and precision (β), are well-known information retrieval measures.

If the overlap passes a fixed coverage threshold ($F_{i,j} \geq t_c$, $t_c = 0.33$), then it is determined that E_i is tracking GT_j .

Configuration

In this context, configuration means the number, the location, and the size of all objects in a frame of the scenario. The result of a tracking approach is considered to be correctly configured if and only if exactly one E_i is tracking each GT_j. To identify all types of errors that may occur, four configuration measures are defined:

- FN - False negative. A GT is not tracked by an E.
- FP - False positive. An E exists and is not tracking a GT.

Identification

In the field of tracking, identification implies the persistent tracking of an GT by a particular E over time. Though several methods to associate identities exist, we adopted an approach based on a "majority rule" 4. A GT_j is said to be identified by the E_i that passes the coverage test for the majority of its lifetime, and similarly E_i is said to identify the GT_j that passes the coverage test for the majority of its lifetime (this implies that associations between GTs and Es will not necessarily match).
In this approach there arise two types of identification failures. The first type (FIT) occurs when E_i suddenly stops tracking GTj and another E_k continues tracking this ground truth. The second error type (FIO) results from swapping the ground truth paths, i.e. E_i initially tracks GT_j and subsequently changes to track GT_k.

- FIT - Falsely identified tracker. Occurs when an E_k that passed the coverage test for GT j is not the identifying tracker, E_i.
- FIO - Falsely identified object. Occurs when a GT_k which passed the coverage test for E_i is not the identifying object, GT_j. Additionally, two purity measures are introduced to evaluate the degree of consistency to associations between Es and GTs.
- OP - Object purity. If GT_j is identified by E_i, then OP is the ratio of frames in which GT_j and E_i passed the coverage test ($n_{i,j}$) to the overall number of frames GT_j exists (n_j).
- TP - Tracker purity. If E_i identifies GT_j , then TP is the ratio of frames in which GT_j and E_i passed the coverage test ($n_{j,i}$) to the overall number of frames E_i .

More detailed description of the evaluation procedure you can find in the last publications [1][4].

Data collection

The meeting room with technical equipment was used to record several videos by using both omni-directional vision system and classical digital video cameras. The recorded meetings are annotated by semi-automated software. This software was developed for annotation purposes and browsing of the results from tracked sequences. The manual annotations contain head positions defined by size and center. The common setup is with four participants. They are sitting around the table as is depicted on the Fig. 2.

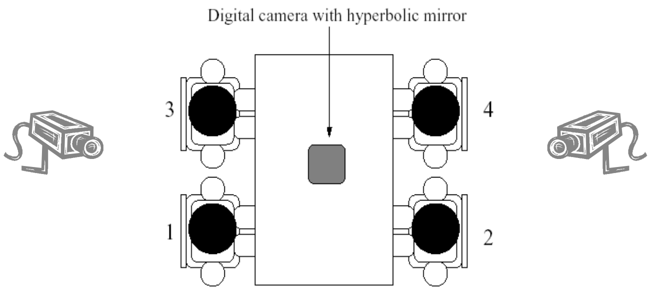

Fig. 2 Setup for participant activity recordings

The omni-directional camera system is placed in the middle of the table. The classical DV cameras were placed on the opposite sides behind the participants. The field of view of these cameras was adjusted to capture participants with part of the table. Each camera captures two participants, who are visible even when they stand. The captured events contain natural meeting with several gestures, etc. Video files from perspective cameras have name left and right, omnidirectional video has name transform and the extension means the correspondence between videos.

Evaluation results

The captured images from video sequences reflect the environment, where the individual meetings were recorded. This chapter contains several sections with captured images from input videos, tables with results for “Face detection” algorithm, KLT algorithm described earlier, and behavior description of various video sources. The results are compared with the type of the video source. The first type of video-sequences called as perspective camera is captured by two classical cameras and the second one called as panoramic is captured by omni-directional system.

Fig. 3 SET1 – omni directional video, left camera, right camera

The first data set was captured with mirror below the camera. The two side cameras observe each two participants from the front. The lighting conditions were affected only by fluorescent lamps. The environment contains background with a lot of nearly skin colored parts, which are mostly visible only in the panoramic images.

Sequence	FN	FP	FIT	FIO	TP	OP
left.a	0,0003	0,0002	0,0000	0,0000	0,9999	0,9997
left.b	0,0001	0,0000	0,0000	0,0000	1,0000	0,9999
right.a	0,0252	0,0288	0,0000	0,3211	0,7667	0,6537
right.b	0,0131	0,0306	0,0000	0,3610	0,8814	0,6259
transform.a	0,5074	0,0153	0,0000	0,1219	0,9775	0,4942
transform.b	0,4878	0,0039	0,0000	0,0002	0,9551	0,6829

Table 1. SET1 - Face Detection results

Sequence	FN	FP	FIT	FIO	TP	OP
left.a	0,0010	0,0473	0,0000	0,0000	1,0000	0,9990
left.b	0,0016	0,0000	0,0000	0,0000	1,0000	0,9984
right.a	0,1849	0,2909	0,0000	0,2652	0,6833	0,6652
right.b	0,1393	0,2845	0,0000	0,4757	0,8693	0,5206
transform.a	0,0892	0,6469	0,0000	0,1190	0,7282	0,7927
transform.b	0,0621	0,6277	0,0000	0,0775	0,8540	0,8607

Table 2. SET1 – KLT results

The sequences from the left camera have the best results because the color balance in the images does not affect so much the skin color detection. The right person on the right camera side has nearly skin colored clothing, which increases FP error mainly for KLT tracker.

Fig. 4 SET2- omnidirectional video, left camera, right camera

The captured omni-directional sequence has worst color properties than images from perspective cameras. The quality of tracking is also affected by objects, which are visible in the omni-directional image and not visible in the partial views from perspective cameras. This becomes evident in the high FP error rate for omni-directional video. The FN error rate is also high for face detection algorithm mainly because of the small resolution of the detected faces. The significant high FIO error rate is evident for right camera sequences, which is again affected by skin color detection failure. Generally, the omnidirectional sequence offers worse results because of the smaller resolution of the captured persons, lighting conditions and cluttered background.

The last data set, which was captured by perspective cameras, was recorded on the different place to achieve various background conditions. The perspective cameras were facing front of participant couple. The field of view is larger and cameras are capturing even the other participants from behind. This setup was chosen because of the miscellaneous environment, where the meetings take place. Such situations and mobility requirements do not enable the best camera positioning in each time.

Sequence	FN	FP	FIT	FIO	TP	OP
left.a	0,0227	0,9724	0,7282	0,0000	0,1287	0,9783
left.b	0,1435	0,3915	0,0000	0,4007	0,9758	0,4563
right.a	0,8779	0,1571	0,0068	0,0444	0,3158	0,0777
right.b	0,8522	0,1645	0,0282	0,1076	0,6182	0,0401
Transform.a	0,3752	0,1635	0,0000	0,3462	0,9565	0,2786
Transform.b	0,3490	0,1941	0,0000	0,3860	0,9617	0,2651

Table 3. SET2 – Face detection results

Sequence	FN	FP	FIT	FIO	TP	OP
left.a	0,9906	0,0024	0,0000	0,0000	1,0000	0,0125
left.b	0,3743	0,1598	0,0000	0,0519	0,9965	0,5740
right.a	0,0781	0,5670	0,0149	0,3153	0,7923	0,6997
right.b	0,1532	0,5107	0,0001	0,4168	0,9118	0,5081
transform.a	0,0108	0,0784	0,0000	0,1331	0,7096	0,8566
transform.b	0,0105	0,0744	0,0000	0,1167	0,8929	0,8729

Table 4. SET2 – KLT results

The situations, where the participants are captured back to the camera, negatively affect the quality of tracking methods. Both the KLT and face detection methods have very big FN error rate in the case of perspective video sequences. The very high FP error rate is caused by cluttered background, which increases the possibility of object misinterpretation. Higher FIT error occurs only in two perspective sequences and frequent tracker terminating by bad image segmentation causes it. The low OP at the panoramic sequences is caused by frequent tracker termination. The TP is still high in this case, which means that the trackers correctly follow given objects. The omni-directional system achieves much better results in all of the measured aspects. The aim of this test was to show that inappropriately positioned perspective cameras could lead to very bad results of tracking and detection methods. The omnidirectional system outperforms the classical approach by reason of only direct participant capturing in this case.

Fig. 5 SET3 – omnidirectional video

The following video sequences were tested without comparison of the perspective cameras. The SET3 was captured with mirror below the camera; the used mirror has uniform vertical resolution. The lighting conditions were quite uniform and the color balance of the video sequence is better when compared with the previous panoramic videos. The lighting in the room is mostly provided by fluorescent lamps with and partly by outer light.

Sequence	FN	FP	FIT	FIO	TP	OP
transform.a	0,0063	0,6291	0,0000	0,1826	0,7681	0,8111
transform.b	0,0324	0,2641	0,0000	0,2127	0,8767	0,7549
transform.c	0,0523	0,3170	0,0000	0,4775	0,7897	0,4702
transform.d	0,0035	0,0080	0,0000	0,0000	0,9982	0,9965
transform.e	0,0160	0,4813	0,0000	0,2867	0,7777	0,6973

Table 5. SET3 – Face detection results

Sequence	FN	FP	FIT	FIO	TP	OP
transform.a	0,0009	0,1718	0,0000	0,0355	0,8340	0,9643
transform.b	0,0150	0,1555	0,0000	0,1147	0,8102	0,8704
transform.c	0,0072	0,2151	0,0000	0,0866	0,8387	0,9096
transform.d	0,0053	0,0000	0,0000	0,0000	1,0000	0,9947
transform.e	0,0040	0,1142	0,0000	0,1298	0,8144	0,8678

Table 6. SET3 – KLT results

Panoramic sequences in this data set have better lighting conditions than in previous cases. The position of participants is also closer to the system, which increases the resolution of detected body parts. The FN error rate is pretty small, which shows that the good lightning conditions and participant placement near the camera system results in high quality detection. The FP error rate is higher at face detection method because of the false face detection on participant's hands. The TP and OP criterions show quite high results, which could be attributed to the coherence of the trackers.

Fig. 6 SET4 – omnidirectional video

The second individual omni-directional video was captured by the same mirror placed above the camera. This setup restricts the overhead environment and enables better hand tracking. The further advantage is in higher resolution in the upper part, where the participant's heads are. The environment and lighting conditions are the same as in the SET2; only the mirror type is replaced.

Sequence	FN	FP	FIT	FIO	TP	OP
transform.a	0,3231	0,0830	0,0000	0,3274	0,9264	0,4657
transform.b	0,3482	0,0551	0,0000	0,2007	0,9267	0,4511
transform.c	0,5047	0,1222	0,0000	0,3177	0,9626	0,3553
transform.d	0,3304	0,0308	0,0000	0,3042	0,9355	0,4871

Table 7. SET4 – Face detection results

Sequence	FN	FP	FIT	FIO	TP	OP
transform.a	0,0127	0,4762	0,0000	0,1376	0,5663	0,8514
transform.b	0,0302	0,4454	0,0000	0,1540	0,7871	0,8168
transform.c	0,0322	0,5158	0,0000	0,0024	0,6369	0,9657
transform.d	0,0361	0,7058	0,0000	0,1457	0,7158	0,8184

Table 8. SET4 – KLT results

The "Face detection" algorithm has higher FN error rate, which is caused by weak hit rate of the face detection. This affects also the low FP error rate. The high TP criterion and low OP criterion also relates to low face detection hit rate. This algorithm is not affected so much by the type of the input video sequence because all evaluation parameters do not vary independently from each other. The second tracking method also embodies good results for FN error rate. The higher FP error rate is caused by misinterpretation of hands with heads.

Stabilization influence

The same tests as these were performed in the previous chapter were applied on the non-stabilized sequences (images transformed without the stabilization algorithm). The idea was to show the necessity of stabilization algorithm[2] both for human presentation and to increase fruitfulness of tracking and detection methods. Two data sets were used for this purpose – SET2 and SET4; both with mirror above the camera, but hyperbolic mirror is the first and mirror with uniform resolution is the second. The following tables contain the results for stabilized sequences represented in white rows and results for sequences, which were not stabilized – grey rows.

Sequence	FN	FP	FIT	FIO	TP	OP
transform.a	0,0105	0,1651	0,0000	0,2013	0,7812	0,7882
transform.b	0,0130	0,1172	0,0000	0,2778	0,7725	0,7093
transform.a	0,0126	0,0845	0,0000	0,1999	0,7368	0,7875
transform.b	0,0129	0,0485	0,0000	0,5027	0,8743	0,4847

Table 9. SET2 - Face detection results(white – stabilized, grey – not stabilized)

Sequence	FN	FP	FIT	FIO	TP	OP
transform.a	0,3231	0,0830	0,0000	0,3274	0,9264	0,4657
transform.b	0,3482	0,0551	0,0000	0,2007	0,9267	0,4511
transform.a	0,3404	0,0540	0,0000	0,1863	0,9878	0,6308
transform.b	0,4633	0,0729	0,0000	0,2771	0,9958	0,3462

Table 10. SET4 – Face detection results(white – stabilized, gray – not stabilized)

The first two tables present the „Face detection" method. The differences are too small to generalize some conclusions. The face detection is done for each frame so it is not so much affected by movement. The correspondence determination was also able to balance the small changes in object positions in order to correctly determine the correspondences between the objects.

Sequence	FN	FP	FIT	FIO	TP	OP
transform.a	0,0277	0,3967	0,0000	0,1159	0,8238	0,8565
transform.b	0,5193	0,2715	0,0000	0,0642	0,8648	0,4165
transform.a	0,0410	0,6245	0,0000	0,2171	0,7943	0,7422
transform.b	0,0712	0,8738	0,0000	0,3369	0,8905	0,5924

Table 11. SET2 - KLT results(white – stabilized, gray – not stabilized)

Sequence	FN	FP	FIT	FIO	TP	OP
transform.a	0,3231	0,0830	0,0000	0,3274	0,9264	0,4657
transform.b	0,3482	0,0551	0,0000	0,2007	0,9267	0,4511
transform.a	0,0117	0,5894	0,0000	0,0367	0,7365	0,9527
transform.b	0,0276	0,5396	0,0000	0,2354	0,6838	0,7372

Table 12. SET4 – KLT results(white – stabilized, gray – not stabilized)

The second tracking algorithm suffers from the vibrations much more than previous detection algorithm. The number of detected objects is higher than when the stabilized sequence is used. The reason is the movement of the whole image, which affects the background subtraction algorithm; therefore, also the image segmentation is affected. This problem decreases the FN error rate, but dramatically increases the FP error rate, which results in a lot of false detections. The CD criterion is also affected by tracking problems.

Conclusion

The most significant advantage of catadioptric systems is in its large field of view with no moving parts. Computer applications can benefit from this advantage and therefore it would be desirable to focus the research into this area. Behavior of the tracking methods was evaluated on various kinds of captured data. The acquired data set contains several video sources from both perspective cameras and omni-directional system. These tests show that the panoramic images acquired from catadioptric systems can be used for human activity monitoring with nearly the same effectiveness as from perspective cameras. However, the conditions of good lightning, high resolution and distortion elimination must be satisfied. The performed tests also showed that artifacts originated by the vibrations can decrease the ability of some tracking and detection algorithms, so it is necessary to use the stabilization algorithms. In addition the performed tests point to some advantages of omni-directional system for human capturing. Inappropriately positioned classical cameras can lead to worse tracking and detection results than these obtained from omni-directional system. The advantage of such system resides in direct to the camera positioning of all people sitting around. The further important advantage is in creating one whole panoramic image that enables better correspondence determination between objects than from images captured by several cameras.

Acknowledgements

This work has been partly supported by the Ministry of Education, Youth and Sports of the Czech Republic under the research program LC-06008 (Center for Computer Graphics).

References

1. Smith, K., Schreiber, S., Potucek, I., Beran, V., Rigoll, G., Gatica-Perez, D., Multi-Person Tracking in Meetings: A Comparative Study, 3rd Joint Workshop on Multimodal Interaction and Related Machine Learning Algorithms, Washington, USA, May 2006.
2. Potúček, I., Automatic Image Stabilization for Omni-Directional Systems, In: Proceedings of the Fifth IASTED International Conference on VISUALIZATION, IMAGING,AND IMAGE PROCESSING, Benidorm, ES, ACTA Press, 2005, p. 338-342.
3. Potúček, I., Sumec, S., Zemčík, P., AUTOMATIC MOBILE MEETING ROOM, In: Proceedings of 3IA'2005 International Conference in Computer Graphics and Artificial Intelligence, Limoges, FR, 2005, p. 171-177, ISBN 2-914256-07-8.
4. Smith, K., Ba, S., Odobez, J., Gatica-Perez, D., Evaluating Multi-Object Tracking, CVPR Workshop on Empirical Evaluation Methods in Computer Vision (EEMCV), San Diego, CA, June 2005.
5. Bunschoten, R., Mapping and Localization from a Panoramic Vision Sensor, Febodruk B.V., Enschede, The Netherlands, ISBN 90-9017279-3, November 2003.
6. Nayar, S., Baker, S., A Theory of Catadioptric Image Formation, Department of Computer Science, Columbia university, Technical Report CUCS-015-97.