Optimizing Camera Placement in Motion Tracking Systems

Dávid Szalóki, Kristóf Csorba, Gábor Tevesz
Department of Automation and Applied Informatics, Budapest University of Technology And Economics, Magyar Tudósok körútja 2/Q., H-1117 Budapest, Hungary

Keywords: Localization Accuracy, Multi-Camera System, Motion Tracking, Optimal Camera Placement.

Abstract: This paper discusses the placement of cameras in order to achieve the highest possible localization accuracy. It is reached by using several cameras with redundant fields of views. A camera model is introduced and the components which cause the localization errors are identified. The localization accuracy measure is defined for one and for multiple cameras too. The problem of adding a new camera to the system in order to improve the accuracy is formulated. The method for finding the optimal placement of this new camera is presented. Some features are enumerated which can be applied for getting an advanced method.

1 INTRODUCTION

Object tracking is nowadays a very popular segment of computer vision. There are application where multiple cameras are used to perform this task. The applications can be separated based on the common field of views of the cameras. The first class contains applications with relatively few overlapping in the field of views. This class contains applications like video surveillance systems where the observed area can be maximized. With these systems the object tracking has to be performed with low accuracy requirements. Only the approximated trajectory is interesting like in (Zhou and Aggarwal, 2006), not its accuracy.

The second class contains applications where the localization accuracy can be helpful. There are a lot of algorithms for the *PnP* problem which stands for Perspective-n-Point problem meaning the estimation of the pose of a calibrated camera from n 3D-to-2D point correspondences. There are several papers about the solution of the *P*3*P*, *P*4*P* or the generalized *PnP* problem. Some of them originate the *PnP* problem in the simpler ones. There are given for this estimation iterative (Oberkampf et al., 1996) and non-iterative (Hesch and Roumeliotis, 2011) (Moreno-Noguer et al., 2007) methods as well. In applications using feature point-based camera tracking (Skrypnyk and Lowe, 2004) (Lepetit and Fua, 2006) dealing with hundreds of noisy feature points is required. Detailed summary about the *PnP* can be found in (Wu and Hu, 2006). All of these generalized methods can be improved if the nature of the noise is known. Thus our first goal is to determine the localization accuracy. The calculated localization accuracy can be used to validate the measured localization accuracy as presented in (Szalóki et al., 2013a).

Another popular topic is the object tracking where a high localization accuracy is required. There are applications where only the object locations are used. If the variation of the localization could be supplied the system performance could be improved. A typical application is the robotic football contest (Käppeler et al., 2010). Here the ball is tracked with stereo cameras. There are competitions where a prebuilt multi-camera system can be used by the teams. With this both the ball and the players can be recognized and tracked. The information about the localization accuracy could be useful for precision positioning and aiming.

In (Ercan et al., 2007) a room is built with multiple fixed cameras. An object is tracked while it can be hidden by some other objects from several cameras. In this case the experimental result could be validated with the theory of localization accuracy.

Assume that we have a camera system containing fixed and movable cameras. We would like to track an object in the world as accurately as possible. The movable cameras can be placed by the system optimally so that the localization accuracy gets it's highest possible value. The position and orientation constraints of the movable cameras are the limitations. We would like to suggest a method for placing the movable cameras in order to get the highest possible localization accuracy.

Szalóki D., Csorba K. and Tevesz G..
Optimizing Camera Placement in Motion Tracking Systems.
DOI: 10.5220/0005012202880295
In *Proceedings of the 11th International Conference on Informatics in Control, Automation and Robotics* (ICINCO-2014), pages 288-295
ISBN: 978-989-758-040-6

At first a camera model has to be formulated. The components which cause the localization inaccuracy can be identified. The localization accuracy has to be defined and a measure has to be chosen. The nature of these has to be examined for better understanding their behavior. As a first step they can be formulated in 2D for one observed point. Later this model can be generalized into 3D and for an observed area or volume instead of one observed point. Finally, some extra features can be included as the field of view, the dynamic constraints or the optimization of more than one cameras together.

Please note that the presented algorithms and methods are part of the Smart Mobile Eyes for Localization (SMEyeL) system which is open-source and is written in C++ using the popular OpenCV (Bradski, 2000) computer vision library. The source code, documentation and all the input data for the presented measurements are available for download from our homepage (SMEyeL, 2013).

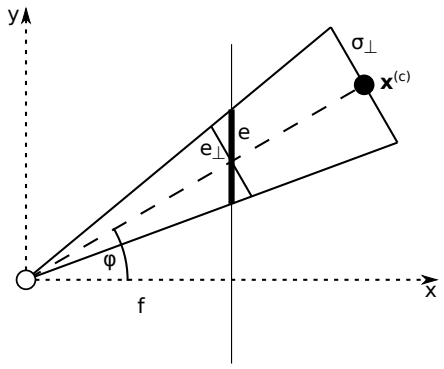

Figure 1: The derivation of the localization error from the detection error. The camera is symbolized with the empty circle. It is facing into the direction x and it has an imager represented by a thin vertical line. f is the focal length. The object is the filled circle which is located at the point $\boldsymbol{x}^{(c)}$ in the camera coordinate system. The object is detected by the camera at φ angle. The detection error e has a perpendicular component $e_\perp$. The perpendicular component of the localization error is $\sigma_\perp$.

2 LOCALIZATION ACCURACY

We assume that we use calibrated cameras. A calibrated camera represents a model, where the distortions are identified and the pinhole camera model is approximated during the mapping. Detailed information about the camera model and the calibration process can be found in (Bradski and Kaehler, 2008). The object localization has a finite accuracy, but then why do we have localization error?

2.1 No Depth Information

One camera by itself can not provide depth information. Based on the image of one camera only a half line can be calculated on which the object is located. Using every half line calculated from every camera can be lead to localize the object. Of course in the real world these half lines do not intersect in a single point. Thus an optimal solution has to be found which can be defined as the object location. Detailed calculations and experimental results can be found in (Szalóki et al., 2013a).

2.2 Detection Error

One camera can detect the tracked object on its own image with an error. This is defined as the detection error. The further the object is, the larger the relative error gets.

The derivation of the localization error in 2D can be followed in Fig. 1. This figure is distorted for the sake of illustration, but in reality the focal length is much larger than the detection error. This means that the following assumption is a good approximation:

$$e_\perp \approx e \cdot \cos\varphi. \tag{1}$$

This error is scaled up to the object location, so:

$$\sigma_\perp \approx d \cdot \cos^2\varphi \cdot \frac{e}{f}, \tag{2}$$

where f is the focal length and e is the detection error of the camera. The distance of the object from the camera is d. φ is the detection angle and $\sigma_\perp$ is the perpendicular component of the localization error. Assuming that the camera is facing towards the tracked object this means that the perpendicular component of the standard deviation is proportional to the distance of the object.

2.3 Covariance Matrix

Therefore, we can calculate the perpendicular component of the localization error of one camera at a single point. As shown in section 2.1 one camera can not provide depth information. This means that the parallel component of the localization error is:

$$\sigma_\parallel \to \infty. \tag{3}$$

For a point in the real world a matrix similar to a covariance matrix can be formulated using (2) and (3). This matrix is diagonal in the coordinate system fitted to the facing direction of the camera and it can

be easily transformed (rotated) into the world coordinate system:

$$\boldsymbol{\Sigma} = \boldsymbol{R}_\alpha^T \boldsymbol{R}_\varphi \begin{bmatrix} \sigma_\parallel^2 & 0 \\ 0 & \sigma_\perp^2 \end{bmatrix} \boldsymbol{R}_\varphi^T \boldsymbol{R}_\alpha, \tag{4}$$

where α is the orientation of the camera and φ is the detection angle. $\boldsymbol{R}_\alpha$ and $\boldsymbol{R}_\varphi$ are the rotation matrices in 2D with angle α and φ respectively. $\sigma_\parallel$ and $\sigma_\perp$ are the parallel and perpendicular components of the localization error.

2.4 Using Multiple Cameras

It has been previously shown that the covariance matrix for one camera observing one single point can be calculated. If more than one camera is used, for each camera its covariance matrix can be determined. These matrices can be combined into one matrix containing the variances of the resulting localization error. This can be achieved by applying the following formulas incrementally:

$$\boldsymbol{\Sigma}_c = \left(\boldsymbol{\Sigma}_1^{-1} + \boldsymbol{\Sigma}_2^{-1}\right)^{-1}, \tag{5}$$

$$\boldsymbol{\mu}_c = \boldsymbol{\Sigma}_c \cdot \left(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_2^{-1}\boldsymbol{\mu}_2\right), \tag{6}$$

where the two original densities are $\mathcal{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ and $\mathcal{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$, while the combined density is $\mathcal{N}(\boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c)$. This is the so called product of gaussian densities.

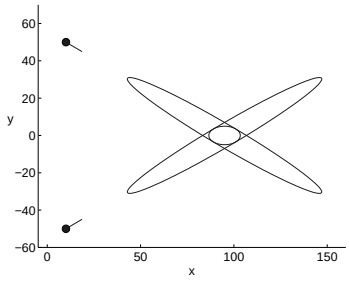

Figure 2: The derivation of the combined covariance ellipse. Two cameras are observing one single point. The cameras are represented with filled circles. The two larger ellipses are the individual covariance ellipses of the cameras separately. For better visualization their major axes are shrunk so that not only two pairs of parallel lines can be seen. The smaller ellipse is the combined covariance ellipse.

A simple example is shown in Fig. 2. The filled circles are the cameras and the short line segments symbolize their directions. The covariance ellipses are drawn at a confidence level of 95 %. The two larger ones are the original ellipses of the cameras. They are distorted, since the limit of their major axis is ∞, but in the figure these axes are chosen in such a way that not only two pairs of parallel lines can be seen. The smaller ellipse is the resulting covariance ellipse. This is not distorted and it can be observed, that the large localization error of one camera disappeared, as the resulting ellipse is smaller.

It can be noticed that the original covariance matrix of one camera contains an element which approaches ∞. Despite this, the inverse of this matrix can be calculated and contains only finite elements. According to (5) the inverse of the final resulting covariance matrix is:

$$\boldsymbol{\Sigma}_r^{-1} = \sum_{i=1}^{n} \boldsymbol{\Sigma}_i^{-1}, \tag{7}$$

where n is the number of the cameras used for localization and $\boldsymbol{\Sigma}_i^{-1}$ is the inverse of the covariance matrix of the i^{th} camera. Since (7) contains only the inverse matrices the inverse of the resulting matrix can be calculated and it contains only finite elements. The determinant of this matrix is zero if and only if there is a specific direction in space in which the localization has a variance with unbounded limit. This can occur if the observed point and the centers of all the cameras fit on the same line. If the determinant of this matrix is non-zero, the resulting covariance matrix can be calculated and the localization has a finite variance in every direction.

In Fig. 3 a camera configuration and the resulting covariance ellipses can be seen. The placement of the ellipses is symmetric, since all the cameras have the same parameters. It can be seen that the errors near

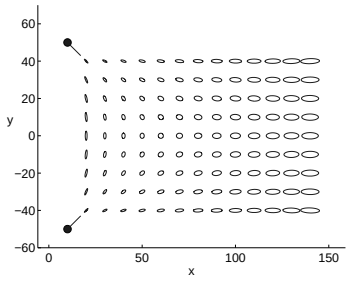

Figure 3: The resulting covariance ellipses of the localization using a specific camera configuration.

the cameras are smaller and they have a larger variance in the y direction, while the distant ones have the larger variance in the x direction.

The next step is to define a way to measure how good the localization is if the tracked object is located in a single observed point. This measure is the localization accuracy and it can be defined in various ways. In this paper two methods are discussed. In both cases the accuracy is derived from the inverse of resulting covariance matrix. The inverse is used because in some cases the inverse exists but the original matrix does not.

2.4.1 Determinant as Measure of Quality

First we can define the accuracy as the determinant of the inverse covariance matrix:

$$q_{det} = \left|\mathbf{\Sigma}_r^{-1}\right|, \tag{8}$$

where q_{det} stands for quality using the determinant. This measure has a physical meaning as it is inversely proportional to the area of the covariance ellipse.

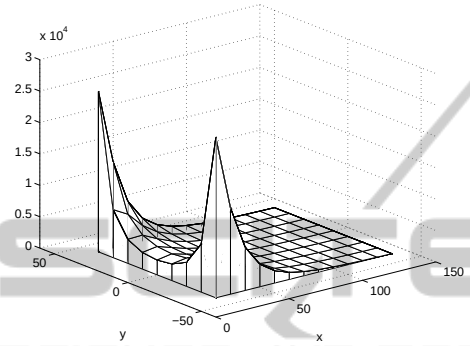

Figure 4: The surface plot of the accuracy in 2D using the previous camera configuration (the same as in Fig. 3) and the determinant measure. The measure is defined in (8).

In Fig. 4 the accuracy is presented in the case of the previous camera configuration. It can be noticed that the closer the observed point is, the better the accuracy becomes. This occurs because the accuracy is proportional to the inverse of the distance (d^{-1}). Of course an object can not be placed arbitrarily close to one of the cameras because a camera has a physical size and minimal focal distance.

2.4.2 Eigenvalue as Measure of Quality

Another possibility is to define the quality as the worst case:

$$q_{eig} = \min\left(\mathrm{eig}\left(\mathbf{\Sigma}_r^{-1}\right)\right), \tag{9}$$

where $\mathbf{\Sigma}_r^{-1}$ is the inverse of the resulting covariance matrix. q_{eig} is the smallest eigenvalue of the inverse matrix. This measure also has a physical meaning as it is proportional to the largest axis of the ellipse.

In Fig. 5 the accuracy is presented in the case of the previous camera configuration using the smallest eigenvalue of the inverse covariance matrix as measure. This is different from Fig. 4, where the determinant was used as measure. Here the best quality is achieved at a point where the detection directions of the two cameras are perpendicular.

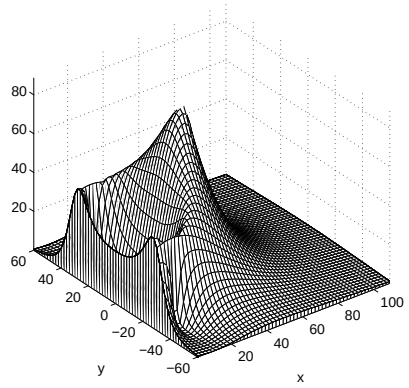

Figure 5: The surface plot of the accuracy in 2D using the previous camera configuration (the same as in Fig. 3) and the eigenvalue measure. The measure is defined in (9).

3 ADD ONE CAMERA

Let there be given a set of fixed cameras to define the camera configuration. These cameras are used for localization at a given point. We would like to add a new camera to the system. The general problem is: how should we place the new camera in order to increase the accuracy as much as possible.
The common method is:

1. Calculate the covariance matrix of the localization determined by the fixed cameras at the observed point.
2. Calculate the covariance matrix parametrized for the new camera at the observed point.
3. Combine these two matrices using (5).
4. Calculate the objective function (q) defined in (8) or in (9).
5. Find the extremal points of this objective function.

Assume that the new camera is placed so that it is facing into the direction of the observed point.

3.1 Using the Determinant Based Measure

In this case the previously described method is used with the determinant measure defined in (8). The localization accuracy in the polar coordinate system fitted to the covariance ellipse generated by the fixed cameras is:

$$q_{det}(\alpha, d) = C + \frac{B + A \cdot \sin^2 \alpha}{d^2}, \tag{10}$$

where α and d are the angle and distance in the polar coordinate system. A, B and C are constant non-negative values which are functions of the camera parameters and the fixed cameras (Szalóki et al., 2013b).

This means that if the new camera is placed at the (α, d) point the localization accuracy at the origin is equal to q_{det}.

This formula can be easily converted into the cartesian coordinate system:

$$q_{det}(x,y) = \frac{A \cdot y^2}{(x^2+y^2)^2} + \frac{B}{x^2+y^2} + C, \qquad (11)$$

where x and y are the coordinates of the new camera. A, B and C are the previously mentioned constant values.

In Fig. 6 the contour of a specific case is plotted. The same two cameras as in Fig. 3 observe the origin and a new camera is added. The curves are "iso-accuracy" curves which means that if the new camera is placed anywhere on a curve the resulting localization accuracy becomes the same value. A curve contained by another one has a higher localization accuracy. It can be noticed from (11) that the plot has reflectional symmetry with the x and y axes as well. In this case the covariance matrix generated from the fixed cameras has two eigenvalues which rate is 1 to 3. If this rate is smaller than $\frac{1}{2}$, the contour plot becomes similar: q_{det} has two local maxima, according to the partial derivative by y. If the rate is larger than $\frac{1}{2}$, the curves become convex ones and q_{det} has only one local maximum using a fixed y. Detailed calculations and experimental results can be found in (Szalóki et al., 2013b).

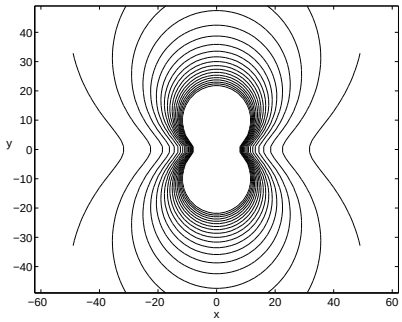

Figure 6: The contours of the localization accuracy with the newly added camera. The covariance matrix of the localization generated by some fixed cameras is known at the origin. A new camera is added to the system. The value of the localization accuracy with the new camera can be calculated. It is a function of the position of the new camera and it has the same value along one curve. The closer the new camera to the origin is, the higher the accuracy becomes.

3.2 Searching on the Boundary

Both of the partial derivatives of (11) equals zero only at the origin, but a camera can not placed right at the observed point. So this means that if a new camera is placed with some constraints, the optimal placement fits on the constraint boundary. This simplifies the problem, since the extremal points of q_{det} have to be searched only on the boundary.

Another consideration can be that in the polar coordinate system for any fixed α_0:

$$q_{det}(\alpha_0, d_1) \geq q_{det}(\alpha_0, d_2) \iff d_1 \leq d_2. \qquad (12)$$

This results in the optimal placement fitting on the constraint boundary since for any (α_0, d_2) point inside the placing area there exists at least one point (α_0, d_1) on the boundary where (12) is satisfied.

Assume that the boundary can be written as a union of functions with one parameter:

$$\bigcup_i \{(x,y) | x = f_i(t), y = g_i(t), t \in [0..1]\} \qquad (13)$$

If the extremal points are searched for on the i^{th} boundary segment, the following has to be solved:

$$t^{(opt)} = \arg\max_{0 \leq t \leq 1} Q_i(t), \qquad (14)$$

where

$$Q_i(t) = q_i(f_i(t), g_i(t)) = q_i(x,y), \qquad (15)$$

which is the localization accuracy on the boundary segment.

Unfortunately, on a boundary segment there can be more than one local maximum even if the segment is a simple line segment. So, global optimizers should be started for every boundary segments. After that the highest local maximum value can be chosen as a global maximum.

Another possibility is to calculate the derivative of $Q_i(t)$ on every boundary segment and find every t value where:

$$\frac{\partial Q_i(t)}{\partial t} = 0. \qquad (16)$$

This can be a complex task, but fortunately, if the boundary segments are line segments, this is reduced to simple solving of third-degree polynomials, which can be performed analytically. In this case the placing area of the new camera is a polygon. This polygon limitation is not too strong since most of the boundaries can be approximated with polygons.

3.3 Using the Eigenvalue Based Measure

Another possibility is to define the localization accuracy as q_{eig} in (9). The characteristic polynomial of the inverse covariance matrix can be formulated. The smaller root of this polynomial is the accuracy. This has to be maximized. In this case, if a new camera is

added in the polar coordinate system, the localization accuracy is:

$$q_{eig}(\alpha, d) = \frac{1}{2} \cdot \left(D + \frac{F}{d^2} - \sqrt{\frac{F^2}{d^4} + E^2 + \frac{2EF}{d^2} \cdot \cos(2\alpha)} \right), \quad (17)$$

where D, E and F are constant non-negative values which are functions of the camera parameters and the fixed cameras. This equation can also be easily transformed into the cartesian coordinate system, but it gets a little long and it is not so interesting. One might assume that if $d \to 0$ then $q_{eig} \to \infty$, but it is not true. If $d \to 0$ the members in the square root with d in the denominator get dominant and they eliminate the member with d in the denominator outside the root. With substitution of the constant values, the inverse of the originally smaller variance (most accurate direction) gets the upper limit. This can be achieved if the facing direction of the new camera is perpendicular to the direction of the most inaccurate direction.

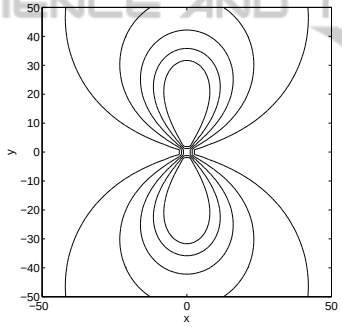

Figure 7: The contours of the localization accuracy with the newly added camera. A curve contained by another one has a higher localization accuracy. The camera configuration and camera parameters are the same as in Fig. 6. The difference is that here the measure defined in (9) is used instead of (8).

So the main difference from the previously discussed q_{det} is that q_{eig} has an upper limit while q_{det} does not. While $q_{det} \to \infty$ if $d \to 0$, q_{eig} has an upper limit. In Fig. 7 the contours of this function can be seen. The camera parameters are the same as in Fig. 6. The basic concept here is also correct that the closer the camera is, the better the localization accuracy gets. However in this case it is more noticable how powerful the angle is. If the new camera is placed parallel to the most inaccurate direction, it improves nothing, while with the determinant measure it does. This behavior represents the real world better than the determinant one.

In this case the optimal placement can also be determined. A consideration similar to (12) can be formulated here as well. This proves that if this line segment and the placing area has no common point it results that the optimum is situating on the boundary. Unfortunately, the boundary can not be handled so simply as in the determinant case because of the square root in (17). This can be handled with global optimization on the line segments. Although it is resource-intensive, it is a simplified method since the optimization has to be performed only on the line segments (1D) instead of the whole area (2D).

3.4 3 Dimensional Case

The previous sections focused on the 2D case. They can be generalized to 3D by increasing the parameter space. The result is that all the positions have three coordinates instead of two. The orientation of the camera is given with three angles rather than one, thus the rotation matrices are the combination of three rotations. There are two independent detection errors and so two independent perpendicular components of the localization accuracy. Thus the covariance matrix is a 3-by-3 matrix instead of a 2-by-2 one. In 3D the objective functions get more complex, specially the measure with the eigenvalue. The optimization can be done only numerically. A consideration similar to (12) can be formulated here as well. This results in the optimum lying on the boundary since the closer point with the same orientation has a larger objective funtcion value. Thus the optimization has to be performed only on the 2D boundary instead of the whole 3D volume.

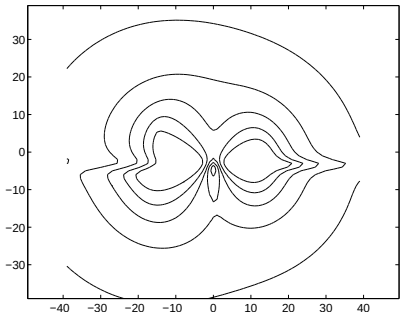

Figure 8: The contours of the objective function defined with the eigenvalue in 3D on a general plane. It has two peaks and a hollow space between them. It means that the angle is powerful and it can radically change the objective function value on a short distance.

In Fig. 8 the contours of the objective function defined with the eigenvalues can be seen on a general plane. It has two peaks and a hollow space between them. It means that there is a direction in which no significant improvement can be achieved but with a little shift in the angle the objective function increases strongly.

4 EXPERIMENTAL RESULTS

We have developed a measurement to confirm the localization accuracy theory presented in Section 2. Three fixed, calibrated cameras observe a marker object. The marker is represented by a printed image that contains easily recognizable ellipses. This marker is moved with a precision robotic arm into each vertex of a 3-by-3-by-3 cubic grid. Tens of localizations are performed at every position on the grid. The mean of these measurements are used as estimations of the positions. The localization accuracy can be characterized by the standard deviation of these estimated marker locations. Detailed calculations, experimental setup and results can be found in (Szalóki et al., 2013a). The standard deviation of marker location estimations can be seen in Fig. 9.

Fig. 10 shows the differences between the standard deviations of the theory and the measurement. Based on our observations, the reprojection error supplied by the OpenCV calibration method is a good choice for estimating the detection error. The measurement differs only with half of a millimeter from the theory in most cases.

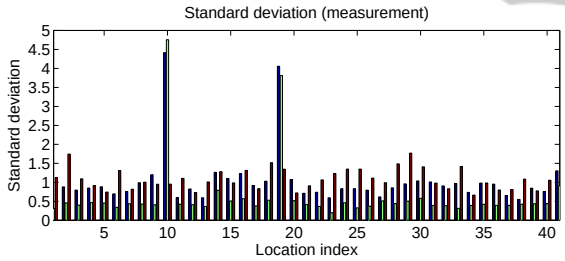

Figure 9: Standard deviation of marker location estimations, shown for the 3 axes separately. Differences are caused by the varying view angles. The two large peaks are measurement errors, they have to be ignored. The direction of the z axis is the nearest to the depth direction. This confirms that the largest inaccuracy can be found in this direction.

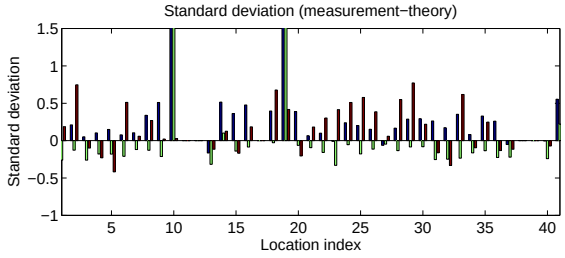

Figure 10: The standard deviation differences of the measurement and the theory in each locations for the 3 axes separately. The two large peaks are measurement errors, as it is mentioned in Fig. 9.

5 FUTURE PLANS

In this section some features are listed which could be applied to the model. Some of them are easy to include but some cases need large changings and developments.

5.1 Area Instead of a Single Point

In the previous sections there is only one observed point where the localization has to be performed. The more general case is that the localization is needed in an area. Therefore a density function can be used. Its value at a specific $\boldsymbol{x}$ point in space represents the probability that the localization has to be done at that point. In this case the objective function can be defined as the expected value of the accuracy:

$$\mathbb{E}_{\boldsymbol{x}}(q(\boldsymbol{x})) = \int_{\Omega} q(\boldsymbol{x})\, f(\boldsymbol{x})\, d\boldsymbol{x}, \tag{18}$$

where Ω represents the whole space and $f(\boldsymbol{x})$ is the density function.

Another method for handling not only one observed point is, when the worst case is used. This means that the localization accuracy on an area is defined as the worst accuracy of its points.

5.2 Taking Field of View into Account

The field of view of the cameras can also be taken into account. This is only a 0-1 function which multiplies its inverse covariance matrix, but it complicates the optimization. An idea is that the space can be cut into pieces with the boundaries of the fields of views. One piece can be handled as described previously and all the pieces has to be managed individually.

5.3 Utilizing Multiple Cameras Together

If multiple cameras can be placed for localization accuracy improvement then applying the previous methods means that the cameras are placed incrementally. The resulting localization accuracy can be increased by optimizing the movable cameras together.

5.4 Rotation in 3D

In 3D if the camera is placed at a fixed position so that it is facing towards the observed point, then the rotation around the facing direction can change the objective function value. This rotation can also have some constraints which have to be taken into account.

5.5 Continuous Optimization for Dynamic Environments

Since a moving object is tracked the optimal placement of the movable cameras should be performed continuously through time. One camera can be placed closer to the object in order to have larger accuracy improvement, but if the object is moving fast it can occur that the camera can not follow it because of its dynamic constraints. It can be a better solution to place the camera further in order to be able to follow the object on a longer path. These dynamic constraints of the movable cameras could be also taken into account.

6 CONCLUSIONS

In this paper the multi-camera localization accuracy and the optimal camera placement is examined. First the camera model is formulated. The localization accuracy is defined for one camera observing one single point. A method for calculating the localization accuracy using multiple cameras is given. All the calculations are performed in 2D and they are extended later into 3D. Two measures are defined. Their benefits and disadvantages are compared. In both cases the objective function is calculated in case of adding a new camera to the system. The optimization of the placement of the new camera is discussed. The general extension into 3D is described. Finally, the future plans are formulated.

ACKNOWLEDGEMENTS

This work was partially supported by the European Union and the European Social Fund through project FuturICT.hu (grant no.: TAMOP-4.2.2.C-11/1/KONV-2012-0013) organized by VIKING Zrt. Balatonfüred.

This work was partially supported by the Hungarian Government, managed by the National Development Agency, and financed by the Research and Technology Innovation Fund (grant no.: KMR_12-1-2012-0441).

REFERENCES

Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools*.

Bradski, G. and Kaehler, A. (2008). *Learning OpenCV*. O'Reilly Media Inc.

Ercan, A., El Gamal, A., and Guibas, L. (2007). Object tracking in the presence of occlusions via a camera network. In *Information Processing in Sensor Networks, 2007. IPSN 2007. 6th International Symposium on*, pages 509–518.

Hesch, J. and Roumeliotis, S. (2011). A direct least-squares (dls) method for pnp. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 383–390.

Käppeler, U.-P., Höferlin, M., and Levi, P. (2010). 3d object localization via stereo vision using an omnidirectional and a perspective camera.

Lepetit, V. and Fua, P. (2006). Keypoint recognition using randomized trees. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(9):1465–1479.

Moreno-Noguer, F., Lepetit, V., and Fua, P. (2007). Accurate non-iterative o(n) solution to the pnp problem. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1–8.

Oberkampf, D., DeMenthon, D. F., and Davis, L. S. (1996). Iterative pose estimation using coplanar feature points. *Comput. Vis. Image Underst.*, 63(3):495–511.

Skrypnyk, I. and Lowe, D. G. (2004). Scene modelling, recognition and tracking with invariant image features. In *Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality*, ISMAR '04, pages 110–119, Washington, DC, USA. IEEE Computer Society.

SMEyeL (2013). Smart Mobile Eyes for Localization (SMEyeL).

Szalóki, D., Koszó, N., Csorba, K., and Tevesz, G. (2013a). Marker localization with a multi-camera system. In *2013 IEEE International Conference on System Science and Engineering (ICSSE)*, pages 135–139.

Szalóki, D., Koszó, N., Csorba, K., and Tevesz, G. (2013b). Optimizing camera placement for localization accuracy. In *14th IEEE International Symposium on Computational Intelligence and Informatics (CINTI)*, pages 207–212.

Wu, Y. and Hu, Z. (2006). Pnp problem revisited. *J. Math. Imaging Vis.*, 24(1):131–141.

Zhou, Q. and Aggarwal, J. (2006). Object tracking in an outdoor environment using fusion of features and cameras. *Image and Vision Computing*, 24(11):1244 – 1255. Performance Evaluation of Tracking and Surveillance.