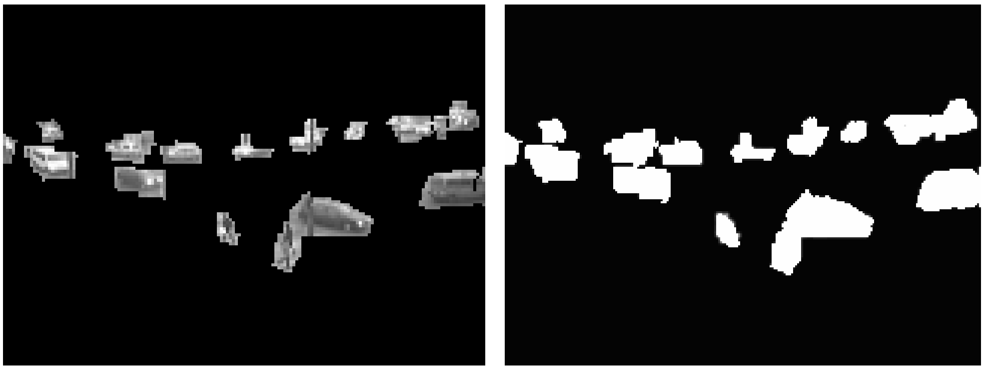
Video camera registration using accumulated motion maps

Zoltan Szlavik, Tamas Sziranyi, Laszlo Havas[i]

Abstract – The paper presents a method to register partially overlapping camera-views of scenes where the objects of interest are in motion even if unstructured environment and motion. In a typical outdoor multi-camera system the observed objects might be very different due to the changes in lighting conditions and different camera positions. Hence, static features such as color, shape, and contours cannot be used for camera registration in these cases. Calculation of co-motion statistics, which is followed by outlier rejection and a nonlinear optimization, does the matching. The described robust algorithm finds point correspondences in two camera views (images) without searching for any objects and without tracking any continuous motion. Real-life outdoor experiments demonstrate the feasibility of our approach.

Keywords –image registration, sequences, multi-sensor system.

Table of Contents:



a) b)

Figure 1. The detected objects are shown in image (a); the corresponding change map is shown in (b).

$$M(I_c,t)=\{m_{ij}(I_c;t)\},\qquad 1\le i\le h,\qquad 1\le j\le w,$$
$$m_{ij}(I_c;t)=\begin{cases}1, & \text{change is detected at pixel } (i,j)\in I_c,\\ 0, & \text{else.}\end{cases}, \tag{2}$$

$$MSM(I_c,p,q)=\frac{1}{\max\limits_{\substack{1\le i\le h\\1\le j\le w}}\left(\sum\limits_{0<t\le T}M(I_c,t)m_{pq}(I_c,t)\right)}\sum_{0<t\le T}M(I_c,t)m_{pq}(I_c,t), \tag{3}$$

where I_c is the input image of the c-th camera; $M(I_c,t)$ is the change map of the t-th frame of the c-th camera, $t\in Z$; $m_{pq}(I_c,t)$ is the value of the motion map at the given pixel $(p,q)\in I_c$ at time t (t-th frame); T is the index of the last frame. Values in $MSM(I_c,p,q)$ assigned to pixel $(p,q)\in I_c$

are the conditional probabilities that change was detected at pixels of I_c when motion was detected at pixel $(p,q) \in I_c$. Examples of such motion statistics, shown as images, can be seen in the left image of Figure 3. Higher the value at a given position of $MSM(I_c, p, q)$ brighter the value of the corresponding pixel in the image.

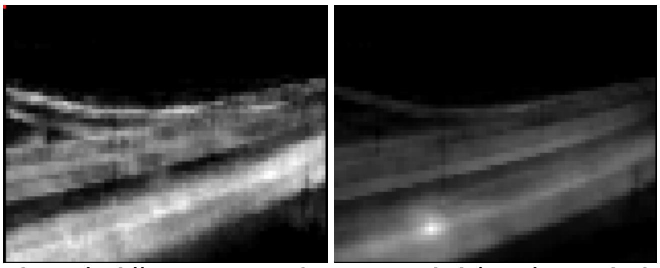


Figure 2**.** Remote statistical maps for different cases are in the pictures. In the left one for a pixel, which is not in the cameras' common field view; in the right one for a pixel from cameras' common field of view.

Local and remote motion statistical maps (LMSM and RMSM respectively) for a given pixel $x = (p,q) \in I_1$ are calculated by the formulas:

$$LMSM(I_1; p,q) = \frac{1}{\max\limits_{\substack{1 \le i \le h \\ 1 \le j \le w}} \left(\sum\limits_{0<t\le T} M(I_1;t) m_{pq}(I_1;t) \right)} \sum_{0<t\le T} M(I_1;t) m_{pq}(I_1;t), \tag{4a}$$

$$RMSM(I_1; p,q) = \frac{1}{\max\limits_{\substack{1 \le i \le h \\ 1 \le j \le w}} \left(\sum\limits_{0<t\le T} M(I_2;t) m_{pq}(I_1;t) \right)} \sum_{0<t\le T} M(I_2;t) m_{pq}(I_1;t), \tag{4b}$$

where $I_c, c \in \{1,2\}$ are the input images of size $h \times w$ (height $\times$ width) pixels; $M(I_c;t)$ is the change map of image I_c at time t (t-th frame); $m_{pq}(I_k;t)$ is a scalar value of the motion map of image I_k at the given pixel (p,q) at time t (t-th frame); t is the index of the current frame and T is the index of the last frame.

Figure 3 Top images: example of co-motion statistics for inlier point-pairs (*LMSM* is on the left and *RMSM* is on the right). Below: the corresponding point-pair is shown in the images of the left and right cameras.

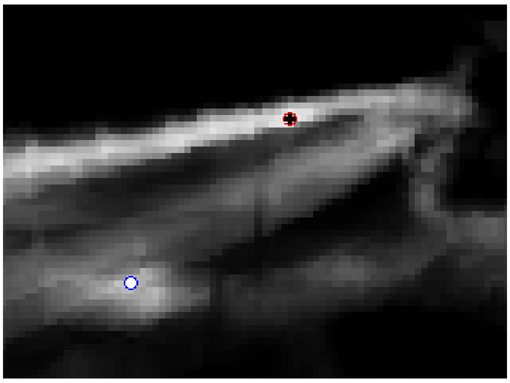


Figure 4. The global maximum is the cross, while it should be the circle.

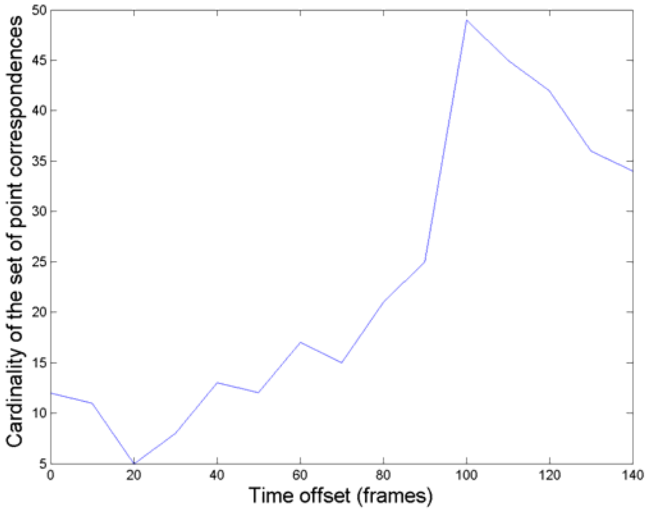


Figure 5. Cardinality of the set of point correspondences for different time offset values. The maximum is at 100 frames, which means that the offset between two sequences is 100 frames.

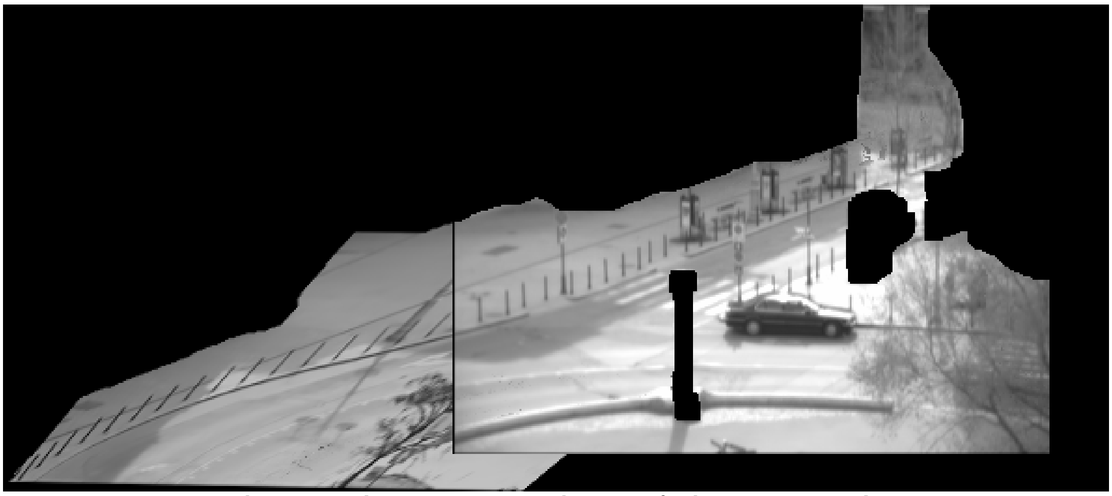


Figure 7. The constructed composite view is in the picture. for the FERENCIEK videos.