placed in front of the VF to compensate for the loss caused by the VF. The spontaneous emission factor for all amplifiers was set to 1.3. In the model, we took into account the saturation effect in the EDFAs by setting the saturation power to 10mW and the relaxation time to 1ms [5]. In this case, periodically stationary BMS propagation proves to be remarkably stable over > 200000km. As shown in Fig. 3, inside the map BMS transmission is bandwidth-guided. In Fig. 3 B_1 = 96.9GHz, B_2 = 111.9GHz for a pulse energy of 410fJ and B_1 = 94.7GHz, B_2 = 80.8GHz for a pulse energy of 171fJ. The filter bandwidths estimated from Fig. 2 are Δ = 142.9GHz and v = 166.7GHz, hence B_{th} = 93.5GHz for the higher pulse energy and B_{th} = 92.1GHz for the lower pulse energy.

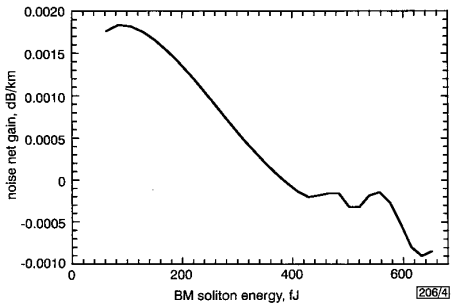

Fig. 4 *Dependence of net gain coefficient for central noise component on energy of BMS*

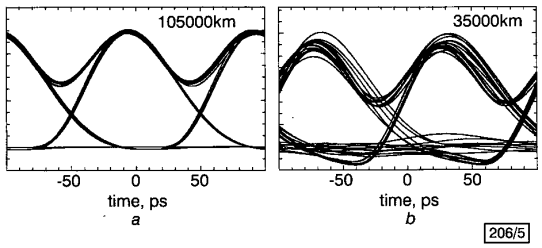

Fig. 5 *Eye diagrams of 64 pseudorandom bit stream transmission with equal number of marks and spaces in system with and without bandwidth management*

a With bandwidth management (BMS) over 105000km
b Without bandwidth management (DMS)
8GHz bandwidth fifth-order Bessel filter is used

Thus, $B_2 > B_{th}$, which means that the criterion in eqn. 2 holds and the noise is discriminated for a pulse energy of 410fJ. Conversely, $B_2 < B_{th}$, $R > 1$ and the noise has a net gain for a pulse energy of 171fJ. In Fig. 4, we calculate the dependence of the net noise gain on the energy of BMS transmission. Note that when the BMS energy is larger than the threshold energy (~380fJ) the noise net gain becomes negative, thus leading to suppression of the noise intensity in the spaces. In Fig. 5, we compare the eye diagrams for the transmission of a 64 pseudorandom bit stream with and without bandwidth management. In the latter case, the parameters of all four amplifiers are the same but the BF and VF filters as well as the 10.5dB compensating amplifier are removed. The remarkable improvement in the eye in Fig. 5*a* compared to Fig. 5*b* is due to three factors. First, the noise in the spaces is almost completely eliminated and does not grow with distance. Note that this suppression occurs despite ~10.5dB extra loss in the case of Fig. 5*a*. Secondly, the amplitude and timing jitter is dramatically reduced due to the strong confinement of the pulse central frequency and bandwidth. Thirdly, the transmission distance is tripled. Future work will focus on analysing the dynamics of BMS transmission depending on the BF and VF parameters as well as the dispersion-managed fibre parameters.

Acknowledgment: This work was supported by an AFOSR grant. The author appreciates the encouragement of G.M. Carter and C.M. Menyuk.

9 February 2000
Electronics Letters Online No: 20000545
DOI: 10.1049/el:20000545

V.S. Grigoryan (*Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County, Baltimore, MD 21228-5398, USA*)

References

1 LEGUEN, D., *et al.*: 'Narrow band 1.02 Tbit/s (51 × 20Gbit/s) soliton DWDM transmission over 1000 km of standard fibre with 100km amplifier spans'. OFC/IOOC'99, San Jose, CA, 1999, Postdeadline Paper PD4

2 FUKUCHI, K., *et al.*: '1.1Tbit/s (55 × 20Gbit/s) dense WDM soliton transmission over 3020km widely dispersion managed transmission line employing 1.55/1.58μm hybrid repeaters'. ECOC'99, Nice, France, 1999, Postdeadline Paper PD2-10

3 MOLLENAUER, L.F., *et al.*: 'The sliding-frequency guiding filter: an improved form of soliton jitter control', *Opt. Lett.*, 1992, **17**, pp. 1575–1577

4 GRIGORYAN, V.S., *et al.*: 'Long distance transmission of filtered dispersion-managed solitons at 40Gb/s bit rate'. Invited talk at ROSC'99 Symp. Massive WDM and TDM Soliton Transmission Systems, Osaka, Japan, 1999. To be published by Kluwer Academic Publishers, 2000

5 MU, R.-M., *et al.*: 'Comparison of theory and experiment for dispersion-managed solitons in a recirculating fiber loop', to be published in *IEEE J. Sel. Topics Quantum Electron.*, 2000, **6**, (1)

Calculation of 3D trajectories of moving objects from unsynchronised stereo video signals

Hyunwoo Kim, Yongduek Seo and Ki Sang Hong

A method for the 3D reconstruction of moving objects from unsynchronised stereo video signals is presented. Using a video mosaicing technique for each video signal, all frames are registered in a reference frame and the trajectories of moving objects obtained as a static structure in a scene, so that correspondence can be determined from an epipolar geometry.

Introduction: In traditional 3D calculation methods, the fact that the stereo sequences are synchronised is typically an essential assumption [1]. However, we have found that 3D information can be extracted without synchronisation when each sequence of stereo video signals is captured by a stationary camera (i.e. without translation). In this case, the trajectories of moving objects can be obtained with respect to a reference frame by using inter-image homographies for image mosaicing. These trajectories can be regarded as a static structure in a scene so that corresponding points can be determined from an epipolar geometry. Obviously, time stamping or synchronisation is not necessary for studying dynamic motion in stereo signals. From the matching results of the trajectories, the projective structure of the moving objects is calculated. Finally, we reconstruct their metric structure, using at least five known points in the background.

Reconstruction method: First, we construct the trajectories of a moving object in a pair of stereo video signals. Assuming that each video image is captured by a stationary camera, video mosaicing techniques calculate the inter-image homographies between a reference frame and other frames in each video signal [2]. Meanwhile, we track the positions of the feature points of a moving object. The tracked positions of each video signal are transferred to its corresponding reference frame and then interpolated using a linear function or splines. As a result, we can obtain a pair of trajectories $\mathbf{s}_l$ and $\mathbf{s}_r$ in the pair of reference frames $\mathbf{I}_l$ and $\mathbf{I}_r$. These trajectories can be regarded as a static structure in a scene so that corresponding points can be determined from an epipolar geometry.

Secondly, we estimate the fundamental matrix $\mathbf{F}$ between the two reference frames, $\mathbf{I}_l$ and $\mathbf{I}_r$, which will be used for the trajectory matching in the following step. Given a minimum of five corresponding points/lines in general positions between the two

frames, the fundamental matrix can be calculated [3]. In addition, the camera projection matrices $\mathbf{P}_l$ and $\mathbf{P}_r$ of two stereo cameras can also be computed up to an arbitrary projective transformation. They are later used for projective reconstruction [4].

Thirdly, the trajectories $\mathbf{s}_l$ and $\mathbf{s}_r$ are matched using an epipolar constraint. When we select a point $\mathbf{p}_l$ in the left trajectory $\mathbf{s}_l$, its corresponding point $\mathbf{p}_r$ in the other view is on the epipolar line $\mathbf{Fp}_l$ and in the right trajectory $\mathbf{s}_r$ at the same time. The crossing points of $\mathbf{Fp}_l$ and $\mathbf{s}_r$ are the candidates of the corresponding point. If we assume that the trajectory of the matching points is smooth, we can easily determine a feasible solution from the candidates by investigating its neighbourhoods. The matching trajectories are represented by $\mathbf{c}_l$ and $\mathbf{c}_r$.

Fourthly, the projective structure of the matching trajectories is computed. Since we determine the corresponding points $\mathbf{c}_l$ and $\mathbf{c}_r$ in the stereo trajectories and the camera projection matrices $\mathbf{P}_l$ and $\mathbf{P}_r$ corresponding to the reference views, the projective structure $\mathbf{X}^P$ of the trajectory of the moving object can be computed using triangulation methods [4]. Note that the projective structure is not meaningful in the Euclidean sense.

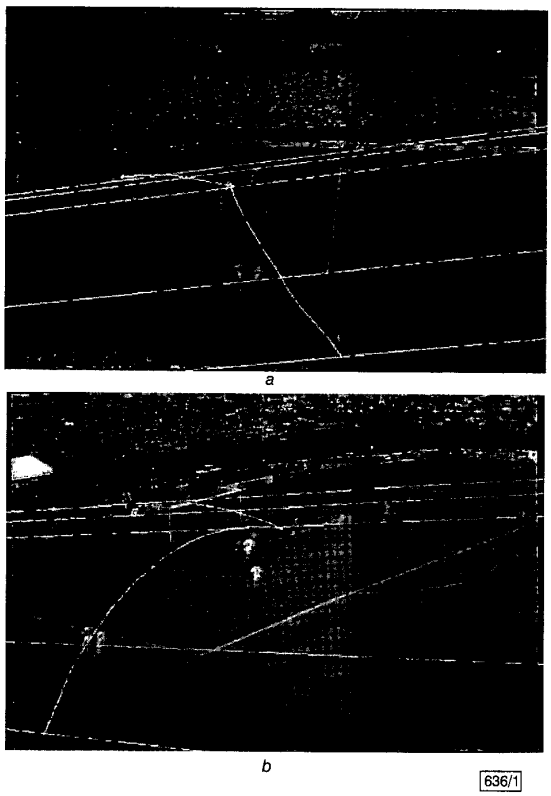

Fig. 1 *Stereo videos and trajectories of moving ball*

a Left reference frame $\mathbf{I}_l$
b Right reference frame $\mathbf{I}_r$

Finally, the projective structure $\mathbf{X}^P$ is upgraded to the Euclidean reconstruction $\mathbf{X}^E$. If we can determine the exact 3D positions of at least five points, then an upgrade is possible. The relationship between $\mathbf{X}^P$ and $\mathbf{X}^E$ is $\mathbf{X}^P = \mathbf{DX}^E$ up to an arbitrary scale factor, where $\mathbf{D}$ is a 4×4 matrix called the projection distortion matrix. If we have a minimum of five known points in the scene, we can compute the projective distortion matrix $\mathbf{D}$. Therefore, the projective structure of the trajectories of moving objects is transformed to a Euclidean structure. Unless we have a sufficient number of known 3D points, then self-calibration methods can be applied [5].

Experiments: We applied our algorithms to a stereo pair of soccer video signals and reconstructed the trajectory of the soccer ball. The sequences were captured from a TV broadcast and they had no time stamp or synchronisation information. Their image mosaics at the reference views $\mathbf{I}_l$ and $\mathbf{I}_r$ are shown in Fig. 1*a* and *b*. The trajectories of the soccer ball were overlaid on them and the trajectories were divided into two sub-trajectories due to the heading of a player. The overlaid lines show the epipolar constraint between the reference frames. Fig. 2 shows the result of the 3D reconstruction. Fig. 2*a* shows a side view and Fig. 2*b* a top view. The square marks are points with known 3D positions, used for estimating the fundamental matrix $\mathbf{F}$ as well as for calculating the projection distortion matrix $\mathbf{D}$. The corresponding positions in the reference frames were selected manually. Errors in the manual selection results in inaccurate 3D reconstruction of those points. That is, the endpoints of the two goal posts, indicated with arrows, should overlap in the top and side views, but the points do not overlap, as seen in Fig. 2*a* and *b*. The standard deviations of the reconstruction error are 10.58, 27.68 and 4.09 (cm) with respect to the *x*-axis, *y*-axis and *z*-axis, respectively.

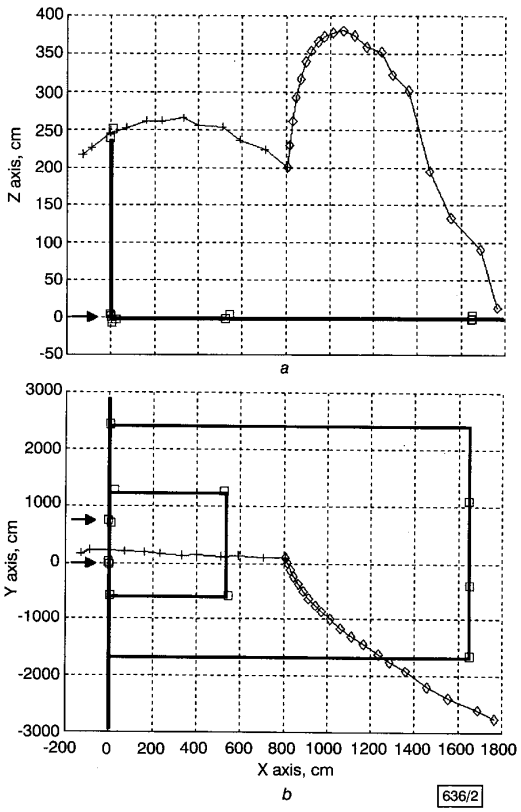

Fig. 2 *3D reconstruction result*

a Side view
b Top view
□ point with known 3D position
◇ first sub-trajectory
+ second sub-trajectory

The first sub-trajectory and the second sub-trajectory seem to be reconstructed reasonably. With these soccer video signals, since an accurate extraction of feature positions is difficult due to motion blurring, the estimated fundamental matrix is inaccurate. Moreover, when the sub-trajectory and the epipolar lines are almost parallel, the matching positions are ambiguous and the reconstruction results become unstable. Fortunately, since we know the start and end positions of the sub-trajectories, we can calculate the correspondences using parameterisation. Each sub-trajectory is parameterised from the start position to the end position and a pair of trajectory positions with the same parameter value is regarded as a correspondence. The reconstruction result is uneven because image registration errors occur due to image blurring. The ball is kicked on the ground and changes direction owing to being headed by a player, with the result that the player scores a goal. In the side view of our reconstruction, the height of the starting position is ~10cm above ground and the ball changes direction at a height of ~200cm. On the goal line, the height of the ball is between the height of the goal post and the ground. We can

guess the headed shot may result in a goal, but its exact determination requires an elaborate quantitative analysis. From the overhead view, we can see that the ball is kicked outside the penalty area and flies with a curved trajectory. Then the ball changes direction due to being headed across the penalty area and passes straight between both goal posts.

Conclusion: We have demonstrated through experiment that our algorithm can reconstruct the 3D structure of moving objects in unsynchronised stereo video sequences. We have demonstrated that 3D information can be extracted without synchronisation when each sequence of stereo video images is captured by a stationary camera.

13 January 2000
Electronics Letters Online No: 20000544
DOI: 10.1049/el:20000544

Hyunwoo Kim, Yongduek Seo and Ki Sang Hong (*Department of Electronics and Electrical Engineering, Pohang University of Science and Technology, Pohang, 790-784, Republic of Korea*)

Ki Sang Hong: Corresponding author

E-mail: hongks@postech.ac.kr

References

1 REID, I., and ZISSERMAN, A.: 'Goal-directed video metrology'. Proc. European Conf. Computer Vision, 1996, Vol. 2, pp. 647–658
2 KIM, H., and HONG, K.S.: 'Robust image mosaicing of soccer videos using self-calibration and line tracking'. POSTECH Technical Report TR-9902, Oct. 1999, Pohang University of Science and Technology
3 FAUGERAS, O.D.: 'What can be seen in three dimensions with uncalibrated stereo rigs?'. Proc. European Conf. Computer Vision, 1992, Vol. 2, pp. 563–578
4 HARTLEY, R.I.: 'Triangulation'. Proc. ARPA IU Workshop, 1994, pp. 190–197
5 POLLEFEYS, M., KOCH, R., and VAN GOOL, L.: 'Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters'. Proc. Int. Conf. Computer Vision, 1998, pp. 90–95

Fractal image compression with simple classification scheme in frequency domain

J.H. Jeng and J.R. Shyu

In fractal image compression, the mean squared error (MSE) computations of the eight orientations of the domain blocks can be reduced into two groups of inner products in the frequency domain. A very simple classification scheme is presented for reducing the two groups of computations into one, which makes the encoder up to 4.7 times faster.

Introduction: Fractal image compression was first developed and implemented by Barnsley and Jacquin [1, 2]. The idea is based on searching for self-similarities. For a given $M \times M$ image, let the range pool be defined as the set of all non-overlapping blocks of size $N \times N$. Here, we assume that N divides M. It is clear that the size of the range pool is $(M/N)^2$. Let the domain pool be defined as the set of all possible blocks of size $2N \times 2N$, which makes up $(M - 2N + 1)^2$ blocks. For simplicity, assume that all the blocks in the domain pool are already reduced to size $N \times N$. For each block in the range pool, the fractal transformation is constructed by searching all the elements in the domain pool to find the most similar block. The similarity allows the dihedral transforms to be made of the blocks, i.e. the eight orientations of the domain blocks generated by rotating the blocks counter-clockwise at angles 0°, 90°, 180° and 270° and flipping with respect to the line $y = x$, respectively.

Given v in the range pool, let u be the search entry in the domain pool. Let u_k, $k = 1, 2, ..., 8$, be the eight orientations of u. Note that $u_1 = u$. The corresponding transforms governing the eight orientations can be represented using the following eight matrices:

$$T_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad T_2 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \quad T_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$
$$T_4 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \quad T_5 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad T_6 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$
$$T_7 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad T_8 = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \tag{1}$$

Thus, eight separate computations of the mean squared error (MSE) are required to search the index k_0 minimising

$$\{\varepsilon_k = \|p_k u_k + q_k - v\|^2 : k = 1, ..., 8\} \tag{2}$$

The quantities p_k and q_k are referred to as the contrast and brightness, respectively, which can be computed directly from the formula

$$p_k = \frac{\left[N^2\langle u_k, v\rangle - \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} u_k(i,j) \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} v(i,j)\right]}{\left[N^2\langle u_k, u_k\rangle - \left(\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} u_k(i,j)\right)^2\right]}$$
$$q_k = \frac{1}{N^2}\left[\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} v(i,j) - p_k \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} u_k(i,j)\right] \tag{3}$$

Similarity in frequency domain: Let $V(m,n)$ be the discrete cosine transform (DCT) of a given image block $v(i,j)$ of size $N \times N$ given by

$$V(m,n) = \frac{2}{N} C(m)C(n) \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} v(i,j) \cos\left(\frac{(2i+1)m\pi}{2N}\right) \times \cos\left(\frac{(2j+1)n\pi}{2N}\right)$$

where

$$c(m) = \begin{cases} \frac{1}{\sqrt{2}} & m = 0 \\ 1 & m \neq 0 \end{cases}$$

Also, let U_k be the DCT of image block u_k. The quantities p_k and q_k in eqn. 3 can be re-formulated as

$$p_k = \langle \hat{u}_k, \hat{v}\rangle / \langle \hat{u}_k, \hat{u}_k\rangle$$
$$q_k = m_v - p m_{u_k}$$

where $m_{uk} = U_k(0,0)$ and $m_v = V(0,0)$ are the mean values, $\hat{u}_k$ and $\hat{v}$ are the de-meaned blocks of u_k and v, respectively. Then each of the eight MSE computations in eqn. 2 can be derived as

$$\begin{aligned} \varepsilon_k &= \langle \hat{v}, \hat{v}\rangle - p\langle \hat{u}_k, \hat{v}\rangle \\ &= \langle \hat{v}, \hat{v}\rangle - \langle \hat{u}_k, \hat{v}\rangle^2 / \langle \hat{u}_k, \hat{u}_k\rangle \\ &= \langle \hat{V}, \hat{V}\rangle - \langle \hat{U}_k, \hat{V}\rangle^2 / \langle \hat{U}_k, \hat{U}_k\rangle \end{aligned} \tag{4}$$

where $\hat{U}_k$ and $\hat{V}$ are the DCT of $\hat{u}_k$ and $\hat{v}$, respectively. Since $<\hat{u}_k, \hat{u}_k> = <\hat{u}_1, \hat{u}_1>$ for all k, then from eqn. 4, finding the index k_0 which minimises eqn. 2 is equivalent to finding the k_0 which maximises

$$\{\tau_k = \langle \hat{U}_k, \hat{V}\rangle, \ k = 1, ..., 8\} \tag{5}$$

The merit of using the frequency domain is to reveal the redundancies in the computations of the eight orientations. As derived in [2], we have

$$\begin{aligned} &\hat{U}_1(m,n) = \hat{U}_1(m,n) \quad \hat{U}_5(m,n) = \hat{U}_1(n,m) \\ &\hat{U}_2(m,n) = (-1)^m \hat{U}_1(m,n) \quad \hat{U}_6(m,n) = (-1)^m \hat{U}_1(n,m) \\ &\hat{U}_3(m,n) = (-1)^n \hat{U}_1(m,n) \quad \hat{U}_7(m,n) = (-1)^n \hat{U}_1(n,m) \\ &\hat{U}_4(m,n) = (-1)^{m+n} \hat{U}_1(m,n) \\ &\hat{U}_8(m,n) = (-1)^{m+n} \hat{U}_1(n,m) \end{aligned} \tag{6}$$

Let $a_{mn} = \hat{U}_1(m,n) \cdot \hat{V}(m,n)$ and $b_{mn} = \hat{U}_1(n,m) \cdot \hat{V}(m,n)$. Thus the eight τ_ks in eqn. 5 are calculated from the following two groups of equations, which can be performed simultaneously: