1

Image/Video Based 3D Modeling, Rendering, and Registration for Virtual Reality

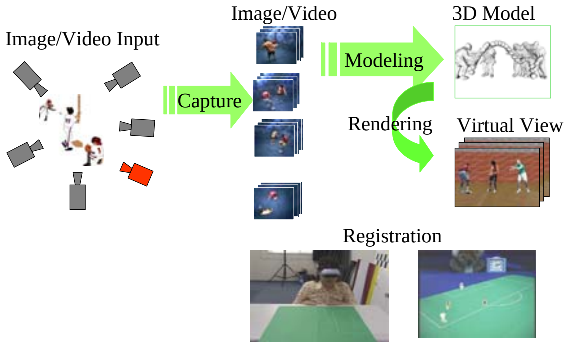
Hideo Saito

saito@ozawa.ics.keio.ac.jp
www.ozawa.ics.keio.ac.jp/Saito

Dept. Information and Computer Science, Keio University, Japan

2

Image/Video Based ...



3

Base Technologies

- Capturing Image
  - Camera Geometry, Calibration
- Modeling and Rendering
  - Shape Recovery
- Registration
  - Camera Calibration
  - Tracking

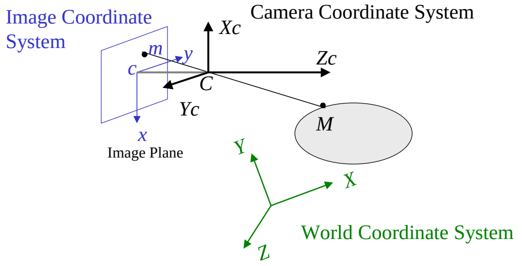
4

Capturing Image

5

Camera Geometry

- Image:Projection of 3D World onto 2D Plane
  - How the projected position is determined ?
  - Relationship between world coordinate, camera coordinate, and image coordinate



6

Camera Coord. and Image Coord.(1/2)

Perspective Projection(Pin-Hole Camera)

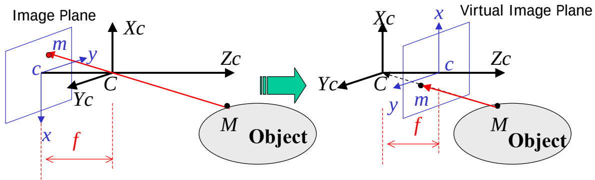


Image Plane is transferred to the front of the lens.

$$x = f\frac{Xc}{Zc}, \qquad y = f\frac{Yc}{Zc}$$

A point on image plane
(x,y) in Image Coord.
(x,y,f) in Camera Coord.

$$\underbrace{s\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}}_{\tilde{\mathbf{m}}} = \underbrace{\begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\mathbf{P}_c} \underbrace{\begin{bmatrix} Xc \\ Yc \\ Zc \\ 1 \end{bmatrix}}_{\tilde{\mathbf{M}}}$$

Projection Matrix

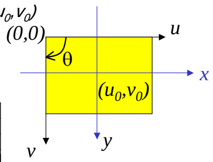
$$s\tilde{\mathbf{m}} = \mathbf{P}_c\tilde{\mathbf{M}}$$

7

Camera Coord. and Image Coord.(2/2)

Digital Image Coordinate

- Horizontal *u-axis*, Vertical *v-axis*.
- Angle between u and v axis θ
- Optical Center (Z-axis of Camera Coord.) (u_0,v_0)
- Scaling $k_u\, k_v$



Projection Matrix $\mathbf{P}_c$:

$$\mathbf{P}_c = \begin{bmatrix} fk_u & -fk_u\cot\theta & u_0 & 0 \\ 0 & fk_v/\sin\theta & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} \alpha_u & -\alpha_u\cot\theta & u_0 & 0 \\ 0 & \alpha_v/\sin\theta & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

⇨ Intrinsic Parameters

$u_0, v_0, fk_u, fk_v, \theta$

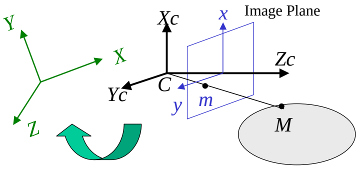
θ=90, k_u=k_v=1 $\quad \mathbf{P}_c = \begin{bmatrix} f & 0 & u_0 & 0 \\ 0 & f & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ ⇨ Intrinsic Parameters u_0, v_0, f

8

World Coord. and Camera Coord.

- Rotation Matrix **R**, Translation Vector **t**

$$\mathbf{M}_c = \mathbf{RM} + \mathbf{t}$$



Position of M

represented in Camera Coord.

$\mathbf{M}_c$=[X_c, Y_c, Z_c]^T

represented in World Coord.

M=[X,Y,Z]^T

(**R**, **t**):Extrinsic Parameters (6DOFs)

$$\begin{bmatrix} Xc \\ Yc \\ Zc \end{bmatrix} = \begin{bmatrix} & 3 & \\ & \mathbf{R} & 3 \\ & & \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} t_X \\ t_Y \\ t_Z \end{bmatrix} \qquad \begin{bmatrix} Xc \\ Yc \\ Zc \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} & \mathbf{R} & & t_X \\ & & & t_Y \\ & & & t_Z \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\mathbf{D}} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$$\mathbf{M}_c = \mathbf{RM} + \mathbf{t} \qquad \tilde{\mathbf{M}}_c = \mathbf{D}\tilde{\mathbf{M}}$$

9

From World Coord. To Image Coord.

- Projection **P:** World Coord. **M** to Image Coord. **m**

$$s\tilde{\mathbf{m}} = \mathbf{P}\tilde{\mathbf{M}}$$

- Camera Coord. $\mathbf{M}_c$ and Image Coord. **m**

$$s\tilde{\mathbf{m}} = \mathbf{P}_c\tilde{\mathbf{M}}_c$$

($\mathbf{P}_c$: Projection Matrix from $\mathbf{M}_c$ to **m**)

- Camera Coord. **M** and World Coord. $\mathbf{M}_w$

$$\tilde{\mathbf{M}}_c = \mathbf{D}\tilde{\mathbf{M}} \qquad s\tilde{\mathbf{m}} = \mathbf{P}_c\mathbf{D}\tilde{\mathbf{M}}$$

P=$\mathbf{P}_c$**D**=**A**[**R**, **t**] $\qquad \theta$=90, k_u=k_v=1

$$= \begin{bmatrix} \alpha_u & -\alpha_u\cot\theta & u_0 \\ 0 & \alpha_v/\sin\theta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_{11} & R_{21} & R_{31} & t_X \\ R_{12} & R_{22} & R_{32} & t_Y \\ R_{13} & R_{23} & R_{33} & t_Z \end{bmatrix} = \begin{bmatrix} f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_{11} & R_{21} & R_{31} & t_X \\ R_{12} & R_{22} & R_{32} & t_Y \\ R_{13} & R_{23} & R_{33} & t_Z \end{bmatrix}$$

P is 11 DOFs

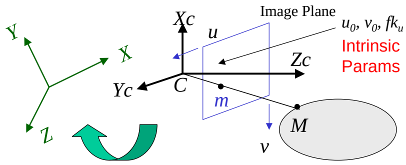
＝Intrinsic 5(3)＋Extrinsic 6

10

Summary of Projection Matrix

- Projection **P:** from World Coord. **M** to Image Coord. **m**

$$s\tilde{\mathbf{m}} = \mathbf{P}\tilde{\mathbf{M}} \qquad s\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{P}\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$



$u_0, v_0, fk_u, fk_v, \theta$

Intrinsic Params

(**R**, **t**):Extrinsic Params(6DOFs)

11

Camera Calibration

- Intrinsic Parameters (5DOFs)
- Extrinsic Parameters (6DOFs)

Basic Method for Camera Calibration

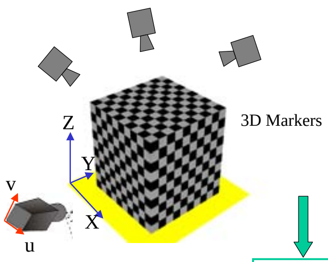
Distribute markers with known 3D positions (X,Y,Z) in objective space

Find image positions (u,v) onto which the markers are projected

$$s\tilde{\mathbf{m}} = \mathbf{P}\tilde{\mathbf{M}} \qquad s\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{P}\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

12

Camera Calibration



Input Data

X1,Y1,Z1 → u1,v1

X2,Y2,Z2 → u2,v2

X3,Y3,Z3 → u3,v3

X4,Y4,Z4 → u5,v5

.

.

Projection Matrix ***P***

Linear Solution for Estimating **P**

$$s\tilde{\mathbf{m}} = \mathbf{P}\tilde{\mathbf{M}} \qquad s\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

(X_n, Y_n, Z_n) --- (u_n, v_n)

$$P_{11}X_n + P_{12}Y_n + P_{13}Z_n + P_{14} - P_{31}X_n u_n - P_{32}Y_n u_n - P_{33}Z_n u_n - P_{34}u_n = 0$$
$$P_{21}X_n + P_{22}Y_n + P_{23}Z_n + P_{24} - P_{31}X_n v_n - P_{32}Y_n v_n - P_{33}Z_n v_n - P_{34}v_n = 0$$

$$\underbrace{\overbrace{\begin{bmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -X_1u_1 & -Y_1u_1 & Z_1u_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -X_1v_1 & -Y_1v_1 & Z_1v_1 \\ & & & & & \cdots & & & & & \\ X_n & Y_n & Z_n & 1 & 0 & 0 & 0 & 0 & -X_nu_n & -Y_nu_n & Z_nu_n \\ 0 & 0 & 0 & 0 & X_n & Y_n & Z_n & 1 & -X_nv_n & -Y_nv_n & Z_nv_n \end{bmatrix}}^{11}}_{2n} \underbrace{\begin{bmatrix} P_{11} \\ P_{12} \\ \vdots \\ P_{32} \\ P_{33} \end{bmatrix}}_{11} = \underbrace{\begin{bmatrix} P_{34}u_1 \\ P_{34}v_1 \\ \vdots \\ P_{34}u_n \\ P_{34}v_n \end{bmatrix}}_{11}$$

n>11/2 markers are required for estimating **P**

Method for extracting intrinsic and extrinsic parameters from projection matrix **P**

$$\mathbf{P=A[R,\ t]} = \begin{bmatrix} \alpha_u & -\alpha_u \cot\theta & u_0 \\ 0 & \alpha_v / \sin\theta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_{11} & R_{21} & R_{31} & t_X \\ R_{12} & R_{22} & R_{32} & t_Y \\ R_{13} & R_{23} & R_{33} & t_Z \end{bmatrix}$$

Projection Matrix **P**→Intrinsic Matrix **A**
Extrinsic Matrix and Vector **R, t**

$$\mathbf{P}_w = \begin{bmatrix} \boxed{P_{11}\ P_{12}\ P_{13}} & \boxed{P_{14}} \\ \boxed{P_{21}\ P_{22}\ P_{23}} & \boxed{P_{24}} \\ \boxed{P_{31}\ P_{32}\ P_{33}} & \boxed{P_{34}} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \\ \mathbf{a}_2 & \mathbf{b} \\ \mathbf{a}_3 & \end{bmatrix} \qquad \mathbf{[R,\ t]} = \begin{bmatrix} \boxed{R_{11}\ R_{21}\ R_{31}} & \boxed{t_X} \\ \boxed{R_{12}\ R_{22}\ R_{32}} & \boxed{t_Y} \\ \boxed{R_{13}\ R_{23}\ R_{33}} & \boxed{t_Z} \end{bmatrix} = \begin{bmatrix} \mathbf{r}_1 & \\ \mathbf{r}_2 & \mathbf{t} \\ \mathbf{r}_3 & \end{bmatrix}$$

Property of Rotation Matrix

$$|\mathbf{r}_1| = |\mathbf{r}_2| = |\mathbf{r}_3| = 1 \quad \mathbf{r}_1 \bullet \mathbf{r}_2 = 0 \quad \mathbf{r}_2 \bullet \mathbf{r}_3 = 0 \quad \mathbf{r}_3 \bullet \mathbf{r}_1 = 0$$
$$\mathbf{r}_1 \times \mathbf{r}_2 = \mathbf{r}_3 \quad \mathbf{r}_2 \times \mathbf{r}_3 = \mathbf{r}_1 \quad \mathbf{r}_3 \times \mathbf{r}_1 = \mathbf{r}_2$$

The property of rotation matrix provide the following solution.

$$\mathbf{r}_3 = \mp \frac{\mathbf{a}_3}{|\mathbf{a}_3|} \qquad \mathbf{r}_1 = \frac{1}{|\mathbf{a}_2 \times \mathbf{a}_3|}(\mathbf{a}_2 \times \mathbf{a}_3) \qquad \mathbf{r}_2 = \mathbf{r}_3 \times \mathbf{r}_1$$

Intrinsic parameter matrix **A** can be obtained as

$$\cos\theta = -\frac{(\mathbf{a}_1 \times \mathbf{a}_3) \bullet (\mathbf{a}_2 \times \mathbf{a}_3)}{|\mathbf{a}_1 \times \mathbf{a}_3||\mathbf{a}_2 \times \mathbf{a}_3|}$$

$$\alpha_u = \frac{|\mathbf{a}_1 \times \mathbf{a}_3|}{|\mathbf{a}_3|^2}\sin\theta, \quad \alpha_v = \frac{|\mathbf{a}_2 \times \mathbf{a}_3|}{|\mathbf{a}_3|^2}\sin\theta$$

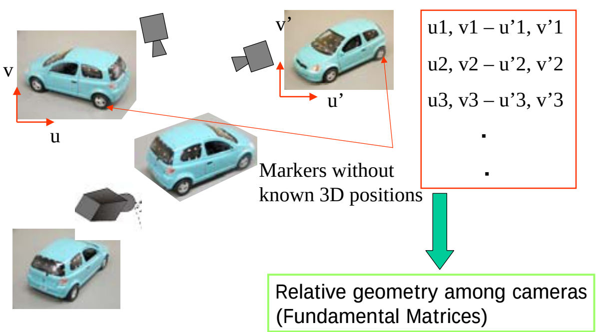
Self Calibration

- Camera calibration without markers

C. Matsunaga and K. Kanatani, “Calibration of a moving camera using a planar pattern: Optimal computation, reliability evaluation and stabilization by the geometric AIC,” Electronics and Communications in Japan, Part 3:Fundamental Electronic Science,Vol. 84, No. 7 pp. 12-21,2001.

M. Pollefeys, R. Koch, and L. V. Gool, “Self-Calibration and Metric Reconstruction in spite of Varying and Unknown Internal Camera Parameters,” International Journal of Computer Vision, 32(1), pp. 7-25, 1999.

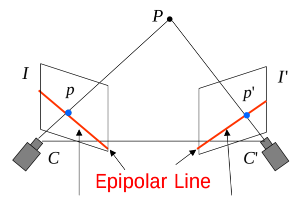
Weak Calibration



Fundamental Matrix

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{pmatrix} \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \mathbf{F}\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} = \mathbf{F}^T \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$



$$ax + by + c = 0 \qquad a'x' + b'y' + c' = 0$$

19

$$\begin{pmatrix}a\\b\\c\end{pmatrix}=\mathbf{F}\begin{pmatrix}x'\\y'\\1\end{pmatrix},\quad \begin{pmatrix}a'\\b'\\c'\end{pmatrix}=\mathbf{F}^T\begin{pmatrix}x\\y\\1\end{pmatrix},\quad (x,y,1)\mathbf{F}\begin{pmatrix}x'\\y'\\1\end{pmatrix}=0$$

One corresponding point gives the following equation.

$$x(x'F_{11}+y'F_{12}+F_{13})+y(x'F_{21}+y'F_{22}+F_{23})+(x'F_{31}+y'F_{32}+F_{33})=0$$

n corresponding points

$$\begin{bmatrix} x_1x_1' & x_1y_1' & x_1 & y_1x_1' & y_1y_1' & y_1 & x_1' & y_1' \\ x_2x_2' & x_2y_2' & x_2 & y_2x_2' & y_2y_2' & y_2 & x_2' & y_2' \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_nx_n' & x_ny_n & x_n & y_nx_n' & y_ny_n' & y_n & x_n' & y_n' \end{bmatrix} \begin{bmatrix}F_{11}\\F_{12}\\F_{13}\\F_{21}\\F_{22}\\F_{23}\\F_{31}\\F_{32}\end{bmatrix} = \begin{bmatrix}1\\1\\1\\1\\1\\1\\1\\1\end{bmatrix}$$

$$F_{33}=1$$

20

Strong/Weak Calibration

	Strong Calibration	Weak Calibration
To be estimated	Projection Matrix $s\begin{bmatrix}u\\v\\1\end{bmatrix}=\mathbf{P}\begin{bmatrix}X\\Y\\Z\\1\end{bmatrix}$	Fundamental Matrix $\begin{bmatrix}a_1\\b_1\\1\end{bmatrix}=\mathbf{F}\begin{bmatrix}x_2\\y_2\\1\end{bmatrix}$
Required data	3D Space – 2D Image Correspondences (Artificial Markers required)	2D Image – 2D Image Correspondences (Natural features)
Size Shape	Available	Unknown (Ambiguity in Projective Transform)

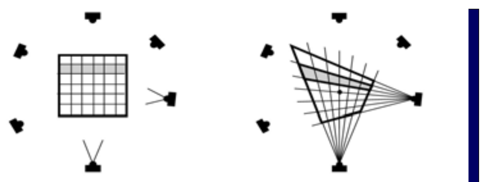
21

Projective Grid Space by Weak Calibration

H. Saito, T. Kanade, Proc. CVPR'99, Vol.2, pp.49–54, 1999.

•Euclid Reconstruction (3D Space is defined independently from cameras)
→**Artificial Maker with Known 3D Position**

•Projective Reconstruction (3D Space is defined dependently on cameras)
→**Natural Maker without 3D Position**



(a)Euclidian Grid Space (b)Projective Grid Space

22

Plane-Based Calibration

A Flexible New Technique for Camera Calibration
http://research.microsoft.com/~zhang/calib/

Multiple Images of Checker Pattern Board at Arbitrary Pose

23

Modeling and Rendering

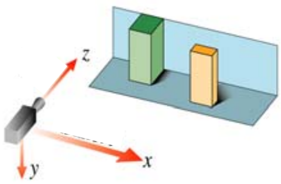
24

Modeling and Rendering

- Image Sequence captured by a camera
  - Motion stereo
  - Feature point tracking
  - Reflectance analysis
- Multiple Cameras
  - Silhouette Intersection
  - Merge Stereo
  - Space Carving/Voxel Coloring

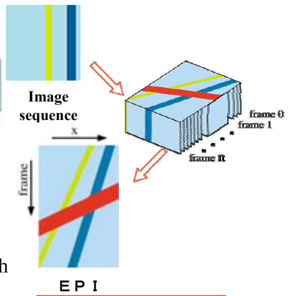
25

Motion Stereo



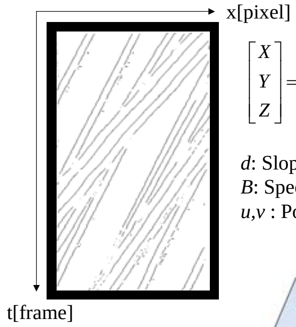
Camera moves toward *x* direction

Epipolar lines are paralell with *x* axis



Epipolar Plane Image

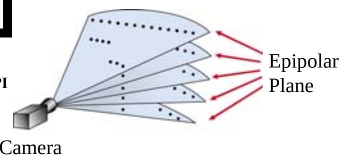
26



Line Detection in EPI

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} uB/d \\ vB/d \\ fB/d \end{bmatrix}$$

d: Slope of line [pixel/frame]
B: Speed of camera [m/frame]
u,*v* : Position of line [pixel]

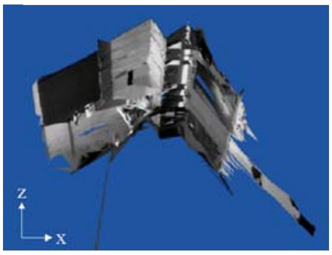


27

120th frame

240th frame

360th frame

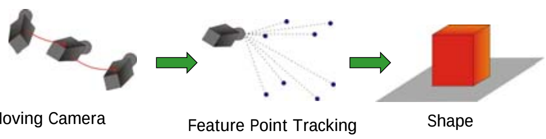


Recovered Shape with Texture of Input Images

28

Feature point tracking

- Feature points are tracked.
- Positions of tracked feature points are used for shape modeling.
- Standard feature tracking algorithm
  - **Kanade-Lucas-Tomasi Feature Tracker (http://robotics.stanford.edu/~birch/klt/)**



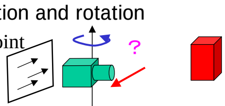
29

Factorization

C. Tomasi and T.Kanade, Shape and Motion from Image Streams under Orthography: a Factorization Method, International Journal of Computer Vision, 9:2, 137-154, 1992

- Estimation of motion of perspective camera
  - Difficult to distinguish translation and rotation

Feature point tracking



  - Depth variation of object shape is relatively smaller than distance to the object.

Orthographic assumption of camera

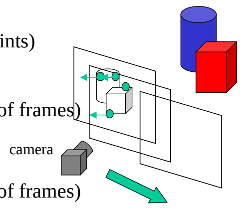
•Easy to distinguish translation and rotation
•Shift of feature points can be segmented into camera motion and object shape via linear computation.

30

Factorization: measurement matrix

P (number of feature points)

$$\tilde{W} = \begin{bmatrix} \tilde u_{11} & \cdots & \tilde u_{1P} \\ \vdots & \tilde u_{fp} & \vdots \\ \tilde u_{F1} & \cdots & \tilde u_{FP} \\ \tilde v_{11} & \cdots & \tilde v_{1P} \\ \vdots & \tilde v_{fp} & \vdots \\ \tilde v_{F1} & \cdots & \tilde v_{FP} \end{bmatrix} \begin{array}{l} \left.\vphantom{\begin{matrix} a \\ a \\ a \end{matrix}}\right\} F \text{ (number of frames)} \\ \left.\vphantom{\begin{matrix} a \\ a \\ a \end{matrix}}\right\} F \text{ (number of frames)} \end{array}$$

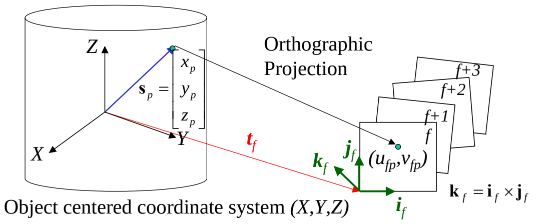


$$\tilde u_{fp} = u_{fp} - a_f \qquad a_f = \frac{1}{P}\sum_{p=1}^{P} u_{fp}$$

$$\tilde v_{fp} = v_{fp} - b_f \qquad b_f = \frac{1}{P}\sum_{p=1}^{P} v_{fp}$$

31

Shape and Motion under Orthographic Projection



$$\tilde{u}_{fp} = u_{fp} - a_f$$

$$u_{fp} = \mathbf{i}_f^T(\mathbf{s}_p - \mathbf{t}_f) \qquad = \mathbf{i}_f^T(\mathbf{s}_p - \mathbf{t}_f) - \frac{1}{P}\sum_{q=1}^{P} \mathbf{i}_f^T(\mathbf{s}_q - \mathbf{t}_f)$$

$$v_{fp} = \mathbf{j}_f^T(\mathbf{s}_p - \mathbf{t}_f) \qquad = \mathbf{i}_f^T\left(\mathbf{s}_p - \frac{1}{P}\sum_{q=1}^{P}\mathbf{s}_q\right) = \mathbf{i}_f^T\mathbf{s}_p$$

32

$$\tilde{u}_{fp} = \mathbf{i}_f^T\mathbf{s}_p$$

$$\tilde{v}_{fp} = \mathbf{j}_f^T\mathbf{s}_p$$

$$\tilde{\mathbf{W}} = \overbrace{\begin{bmatrix} \tilde{u}_{11} & \cdots & \tilde{u}_{1P} \\ \vdots & \tilde{u}_{fp} & \vdots \\ \tilde{u}_{F1} & \cdots & \tilde{u}_{FP} \\ \tilde{v}_{11} & \cdots & \tilde{v}_{1P} \\ \vdots & \tilde{v}_{fp} & \vdots \\ \tilde{v}_{F1} & \cdots & \tilde{v}_{FP} \end{bmatrix}}^{P\text{(number of feature points)}} = \begin{bmatrix} \mathbf{i}_1^T\mathbf{s}_1 & \cdots & \mathbf{i}_1^T\mathbf{s}_P \\ \vdots & \mathbf{i}_f^T\mathbf{s}_p & \vdots \\ \mathbf{i}_F^T\mathbf{s}_1 & \cdots & \mathbf{i}_F^T\mathbf{s}_P \\ \mathbf{j}_1^T\mathbf{s}_1 & \cdots & \mathbf{j}_1^T\mathbf{s}_P \\ \vdots & \mathbf{j}_f^T\mathbf{s}_p & \vdots \\ \mathbf{j}_F^T\mathbf{s}_1 & \cdots & \mathbf{j}_F^T\mathbf{s}_P \end{bmatrix} = \begin{bmatrix} \mathbf{i}_1^T \\ \vdots \\ \mathbf{i}_F^T \\ \mathbf{j}_1^T \\ \vdots \\ \mathbf{j}_F^T \end{bmatrix} \begin{bmatrix} \mathbf{s}_1 & \cdots & \mathbf{s}_P \end{bmatrix} = \mathbf{RS} \quad \Big\} 2F$$

3, 2F, P, 3

Camera Rotation Matrix

Shape Matrix

$\boxed{\text{RANK of } \tilde{\mathbf{W}} \text{ is } 3}$

33

Segmentation of Measurement Matrix into Shape and Motion

$$\tilde{\mathbf{W}} = \mathbf{O}_1\Sigma\mathbf{O}_2 = \underbrace{\begin{bmatrix}\mathbf{O}_1\end{bmatrix}}_{2F \times P}\underbrace{\begin{bmatrix}\Sigma\end{bmatrix}}_{P \times P}\underbrace{\begin{bmatrix}\mathbf{O}_2\end{bmatrix}}_{P \times P}$$

$$\mathbf{O}_1^T\mathbf{O}_1 = \mathbf{O}_2^T\mathbf{O}_2 = \mathbf{I} \quad (P \times P)$$

$$\Sigma = \begin{bmatrix} \sigma_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sigma_P \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{O}'_1 & \mathbf{O}''_1 \end{bmatrix}\begin{bmatrix} \Sigma' & \mathbf{0} \\ \mathbf{0} & \Sigma'' \end{bmatrix}\begin{bmatrix} \mathbf{O}'_2 \\ \mathbf{O}''_2 \end{bmatrix}$$

2F; 3, P-3; 3, P-3; P

$\sigma_1, \sigma_2, \cdots, \sigma_P$: Eigen values of $\tilde{\mathbf{W}}$

$$= \mathbf{O}'_1\Sigma'\mathbf{O}'_2 + \mathbf{O}''_1\Sigma''\mathbf{O}''_2$$

$\boxed{\text{RANK of } \tilde{\mathbf{W}} \text{ is } 3 \Rightarrow \Sigma'' = 0}$

$$\hat{\mathbf{W}} = \underbrace{\begin{bmatrix}\mathbf{O}'_1\end{bmatrix}}_{2F \times 3}\underbrace{\begin{bmatrix}\Sigma'\end{bmatrix}}_{3 \times 3}\underbrace{\begin{bmatrix}\mathbf{O}'_2\end{bmatrix}}_{3 \times P} \qquad \hat{\mathbf{R}} = \begin{bmatrix}\mathbf{O}'_1\end{bmatrix}[\Sigma']^{1/2} \qquad \hat{\mathbf{S}} = [\Sigma']^{1/2}\begin{bmatrix}\mathbf{O}'_2\end{bmatrix}$$

34

Expected factorization

$$\tilde{\mathbf{W}} = \begin{bmatrix} \tilde{u}_{11} & \cdots & \tilde{u}_{1P} \\ \vdots & \tilde{u}_{fp} & \vdots \\ \tilde{u}_{F1} & \cdots & \tilde{u}_{FP} \\ \tilde{v}_{11} & \cdots & \tilde{v}_{1P} \\ \vdots & \tilde{v}_{fp} & \vdots \\ \tilde{v}_{F1} & \cdots & \tilde{v}_{FP} \end{bmatrix} = \begin{bmatrix} \mathbf{i}_1^T\mathbf{s}_1 & \cdots & \mathbf{i}_1^T\mathbf{s}_P \\ \vdots & \mathbf{i}_f^T\mathbf{s}_p & \vdots \\ \mathbf{i}_F^T\mathbf{s}_1 & \cdots & \mathbf{i}_F^T\mathbf{s}_P \\ \mathbf{j}_1^T\mathbf{s}_1 & \cdots & \mathbf{j}_1^T\mathbf{s}_P \\ \vdots & \mathbf{j}_f^T\mathbf{s}_p & \vdots \\ \mathbf{j}_F^T\mathbf{s}_1 & \cdots & \mathbf{j}_F^T\mathbf{s}_P \end{bmatrix} = \begin{bmatrix} \mathbf{i}_1^T \\ \vdots \\ \mathbf{i}_F^T \\ \mathbf{j}_1^T \\ \vdots \\ \mathbf{j}_F^T \end{bmatrix} \begin{bmatrix} \mathbf{s}_1 & \cdots & \mathbf{s}_P \end{bmatrix} = \mathbf{RS}$$

perpendicular

$$\hat{\mathbf{i}}_f^T\mathbf{Q}\mathbf{Q}^T\hat{\mathbf{i}}_f^T = 1$$

$$\hat{\mathbf{j}}_f^T\mathbf{Q}\mathbf{Q}^T\hat{\mathbf{j}}_f^T = 1$$

$$\hat{\mathbf{i}}_f^T\mathbf{Q}\mathbf{Q}^T\hat{\mathbf{j}}_f^T = 0$$

$$\mathbf{R} = \hat{\mathbf{R}}\mathbf{Q}$$

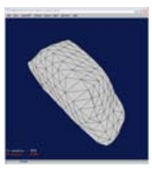
$$\mathbf{S} = \mathbf{Q}^{-1}\hat{\mathbf{S}}$$

$$\hat{\mathbf{R}} = \begin{bmatrix} \hat{\mathbf{i}}_1^T \\ \vdots \\ \hat{\mathbf{i}}_F^T \\ \hat{\mathbf{j}}_1^T \\ \vdots \\ \hat{\mathbf{j}}_F^T \end{bmatrix}, \quad \hat{\mathbf{S}} = \begin{bmatrix} \hat{\mathbf{s}}_1 & \cdots & \hat{\mathbf{s}}_P \end{bmatrix}$$

35

Modeled Shape via Factorization

Input Image Sequence

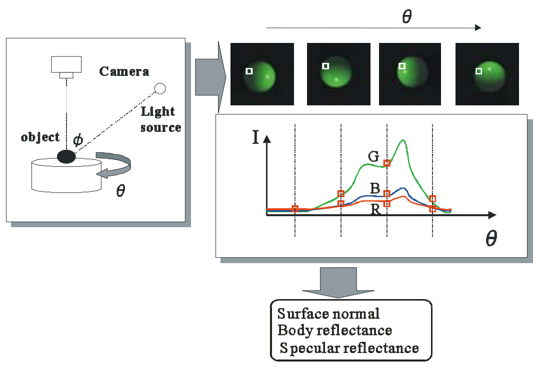


3D Model

Rendered Virtual Views

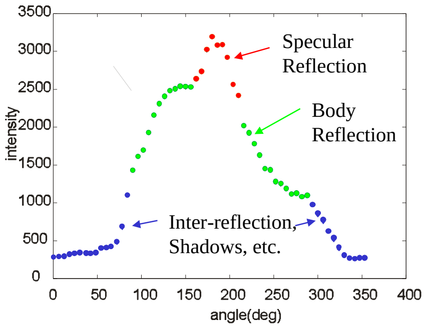
36

Shape Modeling from Reflectance Analysis



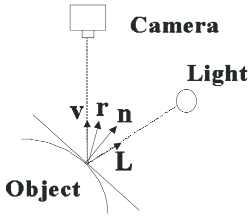
37

Example of Intensity Curve



38

Phong's Reflectance Model



$I = d(\boldsymbol{L}\cdot\boldsymbol{n}) + s(\boldsymbol{r}\cdot\boldsymbol{v})^k$

d: body reflectance

s: specular reflectance

k: specular sharpness

$\boldsymbol{n}$: surface normal

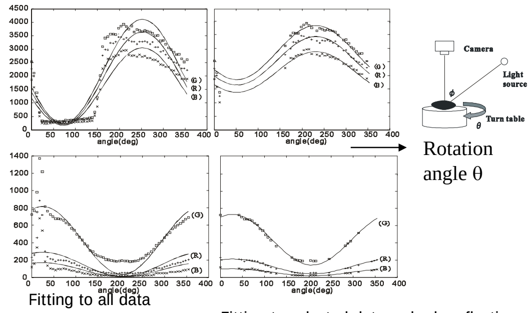
$\boldsymbol{L}$: light source direction

$\boldsymbol{v}$: camera direction

$\boldsymbol{r} = (2\boldsymbol{n}\text{-}\boldsymbol{L})/(|2\boldsymbol{n}\text{-}\boldsymbol{L}|)$

39

Fitting to body reflection

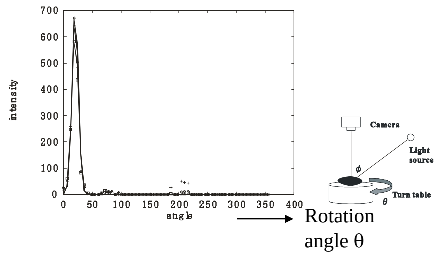


Fitting to all data

Fitting to selected data as body reflection

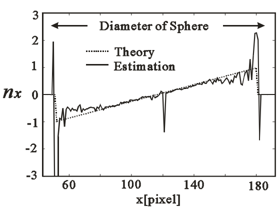
40

Fitting to specular reflection

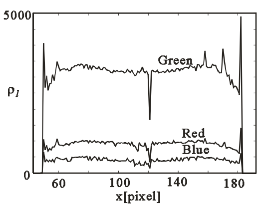


41

Results for Sphere



Comparison of estimated surfaced normal with theoretical value



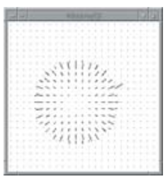
Estimated body reflectance

42

Results for Sphere

Example of input image

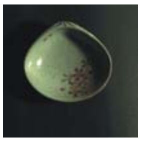
Recovered body reflectance from all intensity data



Recovered needle map

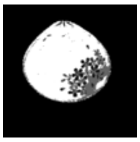
Recovered body reflectance from selected intensity data

Shape and Reflectance

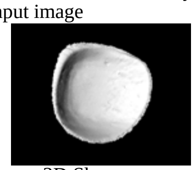


Example of input image

Body reflectance



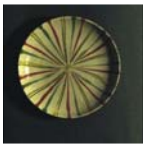
Specular reflectance



3D Shape

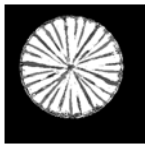
Rendered image at virtual view

Shape and Reflectance

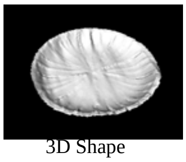


Example of input image

Body reflectance



Specular reflectance



3D Shape

Rendered image at virtual view

Shape and Reflectance

Example of input image

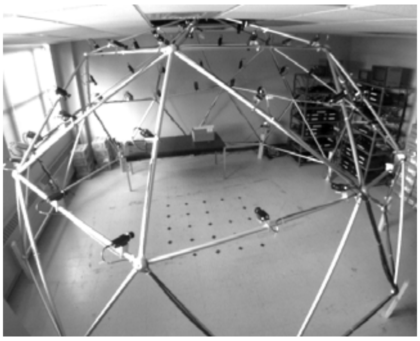
Body reflectance

Specular reflectance

3D Shape

Rendered image at virtual view

Modeling and Rendering via Multiple Cameras



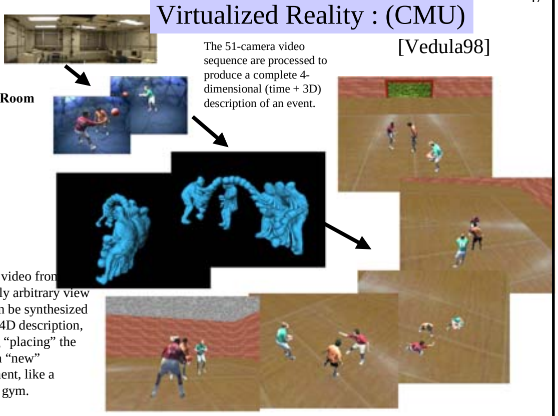
Virtualized Reality: Dome with 51 Cameras in CMU

Virtualized Reality : (CMU)

[Vedula98]

3D Room

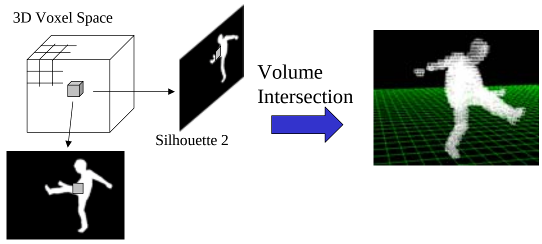
The 51-camera video sequence are processed to produce a complete 4-dimensional (time + 3D) description of an event.



A virtual video from completely arbitrary view points can be synthesized from the 4D description, including "placing" the event in a "new" environment, like a synthetic gym.

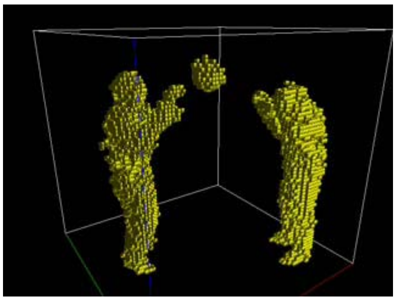
Silhouette Intersection

– Simple, but impossible to recover concaved shape



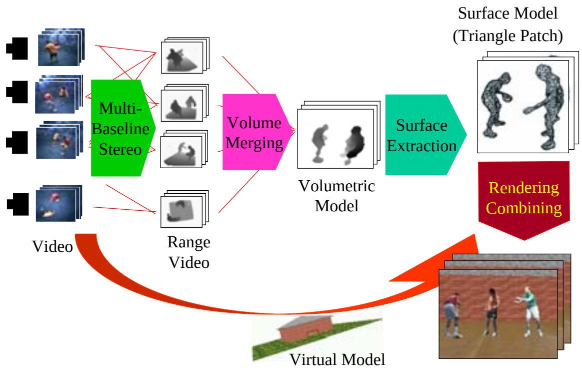
49

Example of Silhouette Intersection



50

Merge Range Images



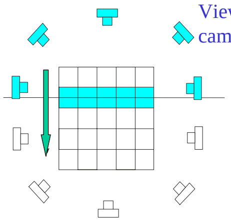
51

Voxel Coloring/Space Curving

Foundations of Image Understanding, L. S. Davis, ed., Kluwer, Boston, 2001, pp. 469-489(Chapter 16, VOLUMETRIC SCENE RECONSTRUCTION FROM MULTIPLE VIEWS/Charles R. Dyer)

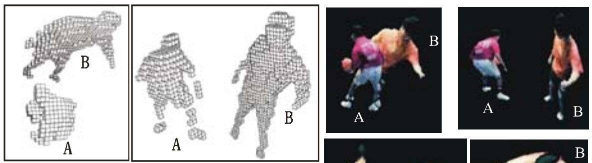
Checking Photo-Consistency of each voxel

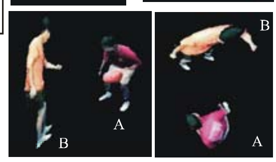
If consistent, the voxel should be on the object surface



52

Shape Modeling and Rendering via Voxel Coloring in Projective Grid Space





- PGS is represented in orthographic grid space.
- Displayed voxel size in (a) is 4 times larger in length than the voxel size in actual shape reconstruction.

53

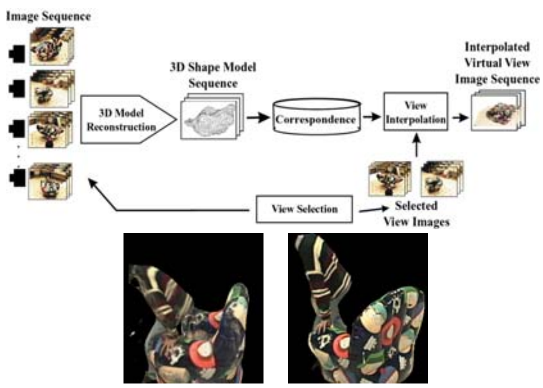
Texture Mapping v.s. View Interpolation

Texture Mapping

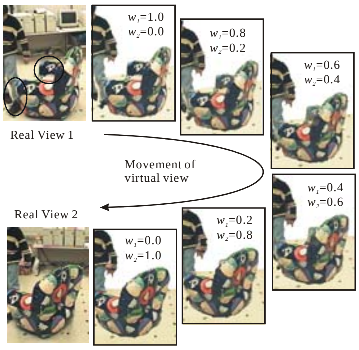
View Interpolation

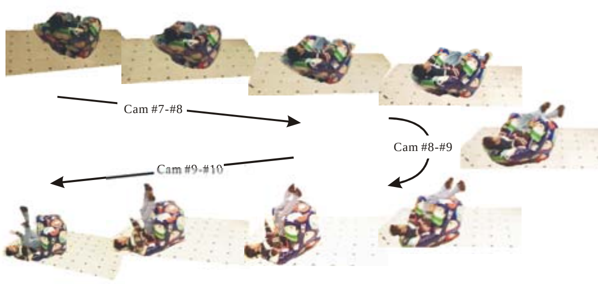
54

View Interpolation via 3D Model



Virtual View Images by Interpolation of Two Views

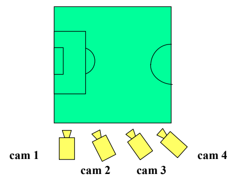


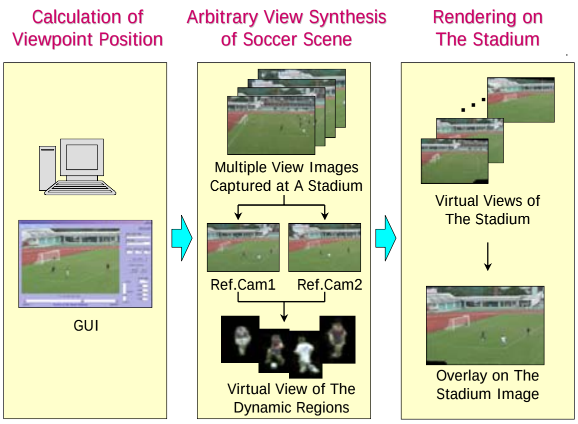


Free Viewpoint Synthesis for Soccer Scene

- Calibration of Multiple Cameras
  Difficult to calibrate the cameras
  – Weak Calibration
- Shape Recovery Techniques
  Almost impossible to recover 3D shapes
  – Simple Shape Representation
- Rendering Methods
  – View Interpolation/Morphing

For soccer scenes

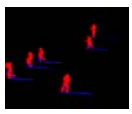




View Synthesis for Dynamic Regions (1/2)

- Detection
  – Subtraction of the background
- Segmentation
  – Player/ball regions
  – Shadow regions

・Color infomation
  HSI transform
・Geometric information
  Homography transform

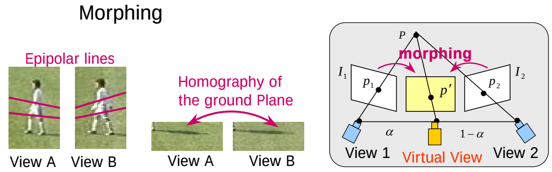


■ :Player/ball regions
■ :Shadow regions

View Synthesis for Dynamic Regions

(2/2)

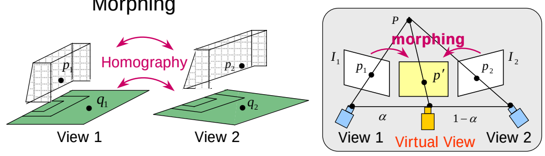
- Player Regions
  - ⇨ Region Correspondence by Homography
    Pixel Correspondence by F-Matrix
    Morphing
- Shadow Regions
  - ⇨ Pixel Correspondence by Homography
    Morphing



62

View Synthesis for Soccer Stadium

- Background Region
  : approximated to plane at infinity
  ⇨ Image mosaicking
- Field Regions
  : approximated to planes
  ⇨ Pixel correspondence by Homography
  Morphing

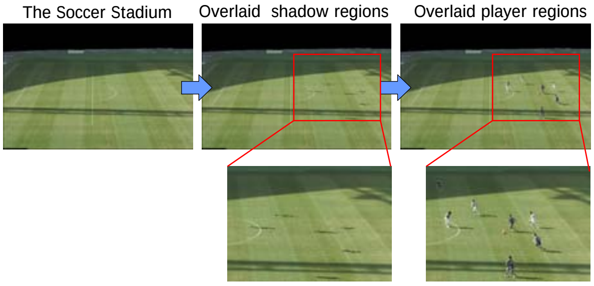


63

Soccer Scene Representation

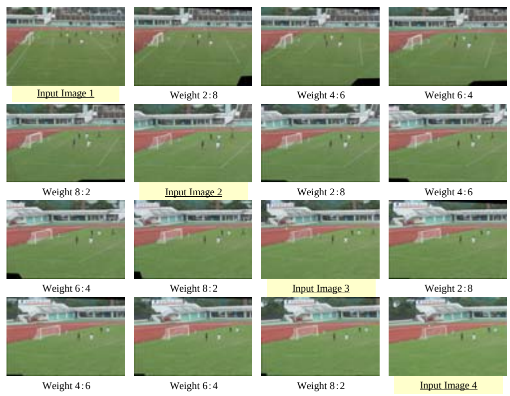
- Superimposing dynamic regions on the soccer stadium.

The Soccer Stadium Overlaid shadow regions Overlaid player regions



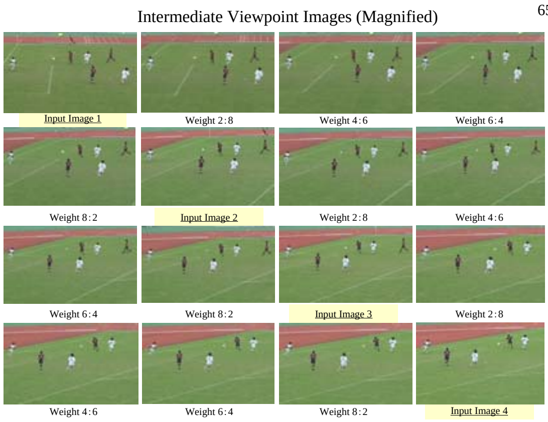
64

Intermediate Viewpoint Images

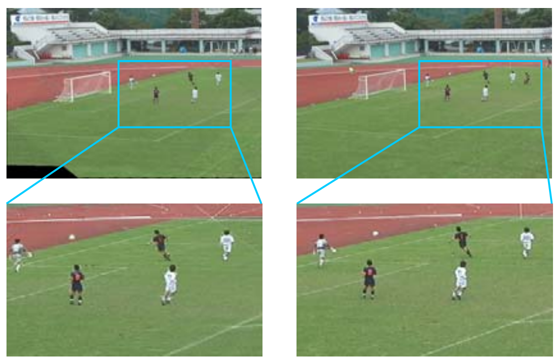


65

Intermediate Viewpoint Images (Magnified)



Comparison



Virtual Camera（ Cam 2-4 Weight 5:5 ）   Real Camera（Cam 3 ）

67

Input Video

Synthesized Video
At Player's Viewpoint

68

Registration

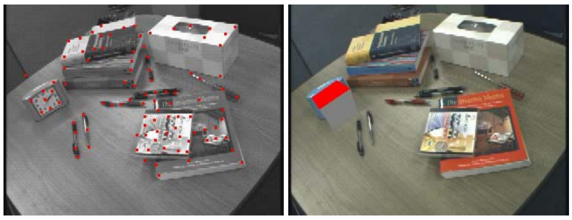
69

Registration for AR/MR

- Feature Tracking is Key Technology.
- Estimate Motion of Camera for Registration.

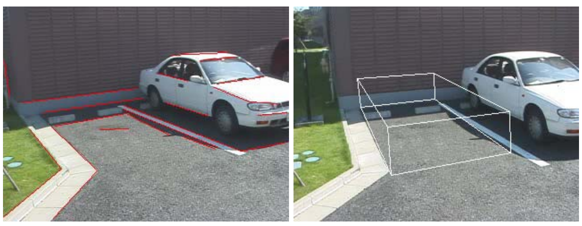
70

Feature point based registration



71

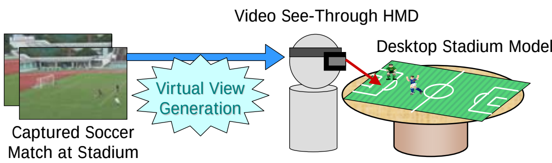
Line-segment Based Registration



72

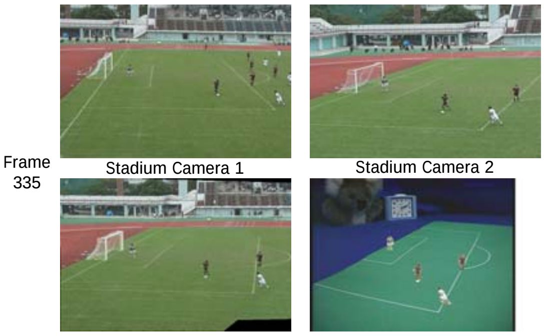
Example of Image Based Registration -Immersive Observation System-

User sees a desktop stadium model in the real world with video see-through HMD and observes dynamic objects of soccer scene overlaid onto the display.



73

Result 1

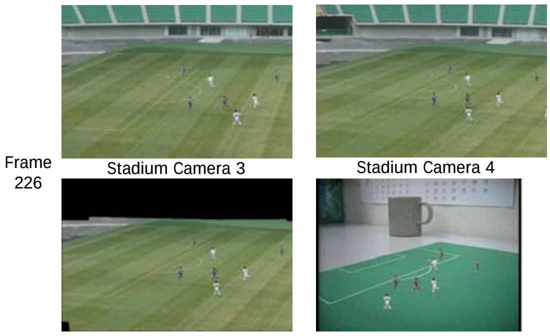


On Real Stadium Image
(Camera 1-2 w = 0.5)

On Tabletop Stadium Model
(Camera 1-2 w = 0.47 z = 1.07)

74

Result 2



On Real Stadium Image
(Camera 3-4 w = 0.9)

On Tabletop Stadium Model
(Camera 3-4 w = 0.87 z = 1.06)

75

Example of Image Based Registration

Texture Overlay onto
Deformable Surface Using HMD

76

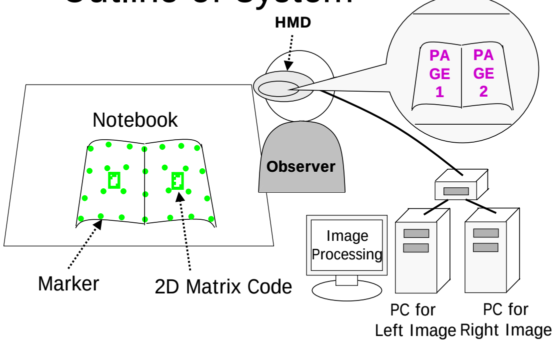
Purpose

Our propose is to overlay textures onto a deformable surface of an object in real time using a video see-through HMD.

Observer feels as if he was reading a real book.

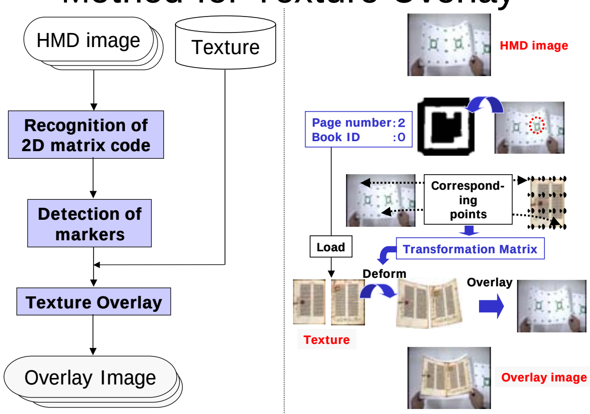
77

Outline of System



78

Method for Texture Overlay



Implementation

79

Input images : 640 × 480 pixels

HMD image

HMD

Appearance of experiment

Texture images : 350 × 500 pixels

book0, page1

book0, page2

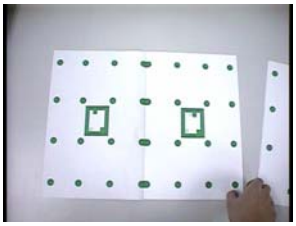
book1, page3

book1, page4

the Gutenberg Bible the Conrad Gesner's Thierbuch

Result 1

80



Original image

Overlay image

Result 2

81

A case that a book is turned upside down.

Overlay image 1

Overlay image 2

Result 3

82

A case that we turn a page.

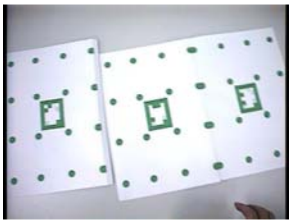
Page 1,2

Page 3,4

Result 4

83

A case that we see multiple books.



Original image

Overlay image

These results(2∼4) shows that 2D matrix code is recognized correctly.

Conclusion

84

- Image/Video-based modeling/rendering, and registration for virtual reality application are introduced.
  - Camera geometry for capturing images
  - Modeling and rendering from image sequence, multiple cameras, reflectance analysis
  - Application of modeling and rendering to sporting scene
  - Application of image-based registration for AR/MR