

## 3차원 다중 기선을 사용한 비데오 영상 모자이크 기술

전재춘 · 서용철

동경대학교 토목공학과

### Video Image Mosaicing Technique Using 3 Dimensional Multi Base Lines

Jaechoon Chon and Yong-cheol Suh

Dept. of Civil Engineering, The University of Tokyo

**Abstract** : In case of using image sequence taken from a moving camera along a road in an urban area, general video mosaicing technique based on a single baseline cannot create 2-D image mosaics. To solve the drawback, this paper proposed a new image mosaicing technique through 3-D multi-baselines that can create image mosaics in 3-D space. The core of the proposed method is that each image frame has a dependent baseline, an equation of first order, calculated by using ground control point (GCP) of optical flows. The proposed algorithm consists of 4 steps: calculation of optical flows using hierarchical strategy, calculation of camera exterior orientation, determination of multi-baselines, and seamless image mosaics. This paper realized and showed the proposed algorithm that can create efficient image mosaics in 3-D space from real image sequence.

**Key Words** : Image mosaics, Image sequence, 3 dimensional multi-baselines, Virtual reality, GIS.

**요약** : 하나의 기선에 작은 여러 영상을 투영하여 하나의 영상 모자이크를 생성하는 2차원 영상 모자이크 기법은 비데오 카메라가 도심지역을 이동하여 얻은 영상을 하나의 모자이크 영상을 생성할 수 없다. 본 논문에서는, 3차원 다중 기준선을 이용함으로써 3차원 공간에서 영상 모자이크를 생성시킬 수 있는 새로운 기법을 제안하였다. 제안된 방법은 각 영상 프레임마다 독립적인 기선을 가지도록 하여 3차원 공간에서 영상 모자이크를 생성 제안하는 것으로서, 독립적인 기선은 각 영상 프레임에서 추출된 광류의 지상 기준점들을 1차 방정식으로 표현한 것이다. 제안한 방법은 계층적 방법을 이용한 광류(Optical Flow)계산, 카메라 외부표정(Exterior Orientation)계산, 다중 기선(Multi-baselines)계산과 모자이크 된 영상들간의 경계를 감지 못하는 화소(Optimal Seamline Detection)의한 영상 모자이크 재생성 과정을 통해 구현되며, 실제 영상 프레임을 이용한 실험을 통해 효과적으로 3차원 공간에서 영상 모자이크 제작이 가능함을 입증하였다.

## 1. 서론

영상 모자이크 기법은 카메라로부터 얻어진 2차원적인 작은 영상들의 중첩 및 연결 조합을 통해 큰 범위의 대형 영상을 제작하는 기술로서, 3차원 영상 데이터 베이스 구축의 경우, 고 해상도 위성영상의 연결 및 확장을 통한 대용량 위성영상의 제작(USGS Hurricane Mitch Program Projects) 및 이를 통한 도시지역에 대한 지형공간정보 취득에 이용되고 있으며, 또한 스테레오 영상 모자이크로부터 다른 시점의 영상(view interpolation) 제작 및 현실적인 3차원 재구성(realistic 3D model), 비디오영상을 이용한 파노라마 영상 구현(Sawhney *et al.*, 1997), 3차원 의료영상 시스템(Chou *et al.*, 1997) 및 가상현실에서의 연속된 화면의 생성(Szeliski, R., 1996) 등에 활발히 적용되고 있다. 특히 최근 GIS분야에서의 영상 모자이크 기법은 3차원 가시화를 통한 안내 및 도로 시설물 등의 효율적이고 체계적인 현황파악 및 유지관리를 가능하게 해 주고 있다.

영상 모자이크 기술은 여러 위성영상을 공간정보 DEM(Digital Element Model)을 이용하여 생성한 정사영상을 공간정보에 매핑하여 영상 모자이크를 제작하는 것과, 공간정보가 없는 일반적인 카메라 영상을 모자이크하는 방법이 있다. 공간데이터에 와핑(warping)기술을 이용하여 텍스처를 매핑하는 기술은 전자기술에 해당된다. 본 논문은 후자에 관한 연구에 속하며, 후자에 관련된 연구 위주로 소개를 한다. 후자의 경우, 일반적인 영상 모자이크 또는 스테레오 영상 모자이크에서 영상에 투영된 물체의 공간정보가 없이 영상 모자이크를 생성하여 비주얼라이제이션 뿐만 아니라, 공간정보 생성, 텍스처 추출 및 공간정보에 텍스처 매핑에 응용할 수 있는 것을 말한다. 영상 모자이크 기술은 카메라의 회전에 의해 얻어진 영상을 실린더 모양의 곡면에 투영하는 파노라마 모자이크(Shum *et al.*, 2000), 카메라 팬/틸트 회전에 의해 얻어진 영상을 반 구형곡면에 투영하는 반구형 모

자이크(Coorg *et al.*, 2000), 카메라 이동에 의해 얻어진 영상을 Fig. 1(a)와 같이 하나의 기선(Baseline)에 투영하여 넓은 범위로 확대된 대형영상을 만드는 2차원 영상 모자이크(Zhu *et al.*, 2001)로 나누어진다. 그러나, 파노라마 모자이크, 반구형 모자이크된 영상에서 공간데이터와 영상 데이터를 취득하고자 할 경우, 한 개의 영상 이미지에서 모든 도시내의 건물, 도로 시설물을 한꺼번에 표현하기는 곤란하므로 여러 영상 이미지를 통해 데이터를 취득한 후, 각각을 연결하여 통합해 주는 과정을 필요로 하게 된다(Coorg *et al.*, 1999). 또한, 카메라 이동을 통해 하나의 기선(Baseline)에 투영하여 생성된 2차원 영상 모자이크를 이용하여 3차원적인 공간데이터와 영상데이터 취득하고자 할 경우에는 각 영상을 연결해 주는 통합과정은 필요치 않으나, Fig. 1(b)와 같이 카메라가 도심지역의 도로를 따라 움직일 때, 하나의 기선으로 영상 모자이크를 생성할 수 없다.

본 논문에서는 이러한 문제점을 해결하기 위해 Fig. 1(b)와 같이 카메라가 도심지역의 도로를 따라 움직일 때, 3차원 공간에서 영상 모자이크를 생성할 수 있는 새로운 기술을 제안하였다. 제안한 방법은 각 영상 프레임마다 독립적인 기선을 가지도록 하여 3차원 공간에서 영상 모자이크 생성하는 것이다. 독립적인 기선은 각 영상 프레임에서 추출된 광류의 지상 기준점들을 1차 방정식으로 표현된 것으로 써, Fig. 1(b)와 같은 경우에도 영상 모자이크가 가능하도록 하였다. 본 연구에서 제안한 방법의 알고리즘은 계층적 방법(Hierarchical Strategy)을 이용한 영상 프레임 간의 특징점 추출 및 정합에 의한 광류(Optical Flow)계산, 광류를 이용한 카메라 외부표정(Exterior Orientation)계산, 그리고 광류들의 지상 기준점을 1차 방정식으로 표현한 다중기선(Multi-baseline)계산, 모자이크된 영상들간의 경계를 감지 못하는 화소(Optimal Seam-line Detection)에 의한 영상 모자이크 재구성을 통해 구현하고 있다.

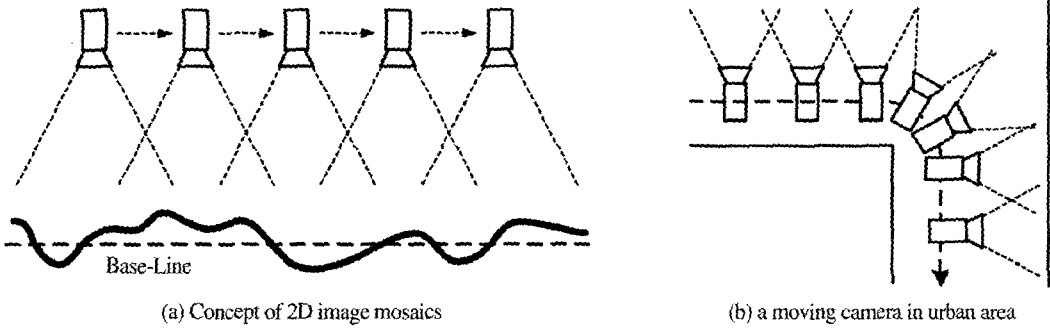


Fig. 1. Conventional image mosaics method.

## 2. 계층적 방법을 이용한 광류(Optical Flow)

광류는 연속된 두 영상프레임간에 동일한 특징점을 연결한 벡터를 의미한다. 광류를 이용하여 두 영상간의 카메라 외부표정을 구할 수 있다. 처음 두 영상간의 카메라 외부표정은 5개의 광류를 이용하여 구할 수 있으며, 3번째 영상 프레임부터 마지막 영상 프레임까지는 광류의 지상기준점을 이용하여 카메라 외부표정(exterior orientation)을 구한다. 광류의 지상기준점은 이전 두 영상 프레임에서 카메라 외부표정을 이용하여 구할 수 있다. 따라서 최소 3개의 광류가 연속된 3개의 영상 프레임동안 추적이 되어야 한다.

광류 3개만을 추적하게 되면 카메라 외부표정 계산시 에러를 많이 포함문제가 있다. 많은 광류가 영상에서 고루 분포되어 있으면 카메라 외부표

정의 에러를 줄일 수 있으므로 본 논문에서는 연속된 3개 프레임에서 많은 광류를 추적하지 못하는 특징점 추적대신 많은 양의 광류를 추적할 수 있는 윤곽선 추적방법 (Contour Tracking)을 사용하였다.

Fig. 2와 같이 원 영상에서 영상 피라미드를 제작하며, 제일 작은 영상 피라미드 크기는  $320 \times 240$  이하이다. Fig. 2와 같이 3개의 영상 피라미드가 있다면, C 알고리즘에서 영상의 대략적인 움직임을 계산하여 B 알고리즘에서 동일한 광류를 찾는 범위를 줄여 계산시간을 줄이게 한다. 또한 B에서 영상의 대략적인 움직임을 계산하여 A 알고리즘에서 동일한 광류를 찾는 범위를 줄여 계산시간을 줄이게 한다. 영상의 대략적인 이동을 구하는 B와 C알고리즘은 이전 영상 프레임과 계산 당시의 영상프레임에서 SUSAN(Smallest Univalve Segment Assimilating Nucleus)(Smith *et al.*, 1997)을 이용하

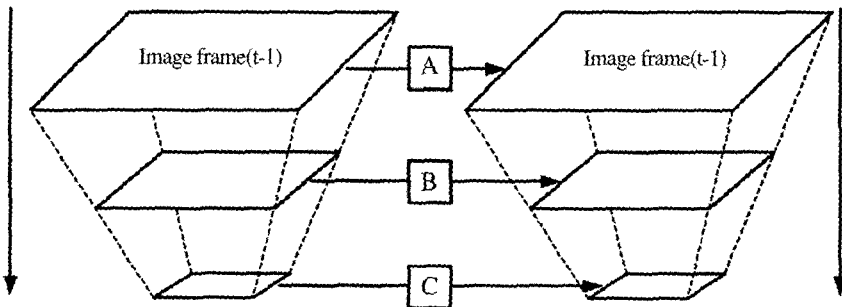


Fig. 2. Fast matching of feature points using hierarchical strategy (t: the number of image frame).

여 특징점을 추출하고 동일한 특징점을 찾아서, 광류들의 평균길이를 구한 것이다.

Smith는 영상에서 에지, 코너(corner)대한 SUSAN원리로 접근하여 사용하였다. 에지 강도 계산은 원형 마스크 중심 (Fig. 3에서 "+") Univalue Segment Assimilating (USAN) 의 영상 값에 원형 마스크 영역 내에 있는 각 영상 값과 차이를 합산 결과 값이다. 이 결과 값을 가우시안 (Gaussian) 함수를 적용하여 나타낸 것이 SUSAN 에지 값이다.

$$S' = \{ \sum e^{-\left( \frac{I(x,y) - I(x+dx, y+dy)}{T} \right)^6} \} / G \quad (1)$$

여기서, I는 화소(gray) 값이고 T는 화소 값의 차에 대한 문턱치(threshold) 값이다. 식(1)에서 T가 크면 화소 값 변화가 큰 곳만을 특징 점으로 추출되고 T가 작으면 화소 변화가 작은 곳에서도 특징 점을 추출한다. 특징 점 선택 대상을 선택하는

기준은 G 값에 달려있다. 일정한 개수의 특징 점들을 구하기 위해서는 T 값을 고정 시켜 지수함수 부분을 계산 후 G 값을 변화시키면 된다. 본 논문에서는 원형 마스크 지름을 크기를 7 화소, T는 30, G는 가변 하도록 설정 하였다. Fig. 3(a)는 7가지의 원형 마스크를 보여주고 있다. 7가지 형태에서 코너점을 뜻하는 a형의 Smallest Univalue Segment Assimilating Nucleus (SUSAN) 값이 가장 작은 값을 가지게 된다. Fig. 3(b)는 원형 마스크의 중심(USAN)이 흰색인 경우이다. 에지 강도에서 특징점을 추출하기 위하여 지역최대(Local Max)를 사용했으며, 추출된 특징점수에 따라 G값이 변하게 된다. Fig. 4는 SUSAN 알고리즘을 적용했을 때 결과를 보여주고 있다.

특징점 정합방법은 상관정합과 SUSAN 에지강도 S'를 융합하여 구하였다(Chon et al, 2001). 상관정합은 영상에서 빠른 정합을 하기 위하여 Sum

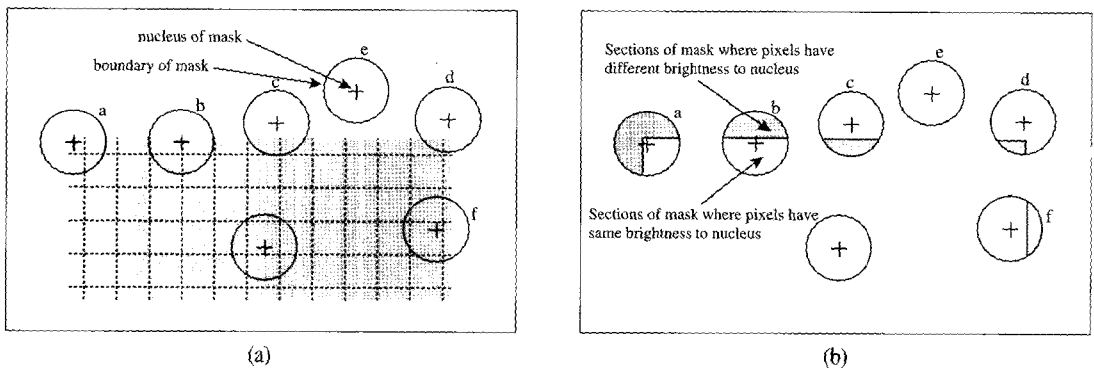
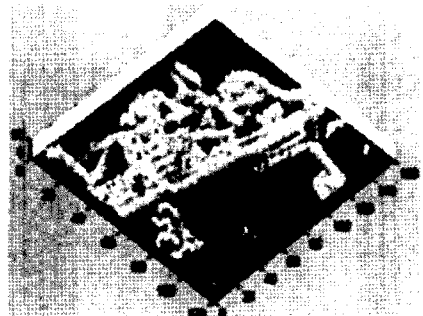


Fig. 3. SUSAN principle: (a) circular masks at different places on the image, (b) USAN show in white color.



(a)



(b)

Fig. 4. Example of SUSAN operation, (a) input image, (b) result by SUSAN.

of Absolute Differences(SAD), Sum of Squared Difference (SAD),과 Sum of Squared Difference (SSD)을 사용한다.

$$SADG = \sum_{i,j=-N/2}^{N/2} |I_1(x+i,y+j) - I_2(x+i+d_x,y+j+d_y)| \quad (2)$$

$$SSDG = \sum_{i,j=-N/2}^{N/2} (I_1(x+i,y+j) - I_2(x+i+d_x,y+j+d_y))^2$$

여기서,  $I_1$ ,  $I_2$ 는 현재와 이전 영상 프레임,  $N$ 은 마스크 크기이다. 상관정합의 그 외 방법으로는 Cross Correlation(CC), Normalized Cross Correlation (NCC)과 Zero-mean Normalized Cross Correlation (ZNCC)방법 등이 있으나, 계산 소요 시간이 많이 들기 때문에, 본 논문에서는 SSD 방법을 사용하였다.

동일한 윤곽선을 찾는 방법은 각 원 영상에서 canny 알고리즘을 적용하여 윤곽선을 추출하고, 영상의 평균이동량을 더한 주변에서 에지 모양과 상관정합으로 윤곽선을 찾는다(Han *et al.*, 2000)

(Fig. 5(a) 참조). 카메라 외부표정을 구하기 위한 광류선정은 영상 프레임을 바둑판 모양의  $5 \times 5$ 개로 분리하여 각 부분에서 가장 정합정도가 높은 광류를 Fig. 5(b)와 같이 추출한다. 선택된 25개의 광류는 정합정도가 높은 순으로 배열하여 카메라 외부표정계산에 사용된다.

### 3. 카메라 외부표정

영상 모자이크를 생성하기 위한 각 영상 프레임의 기선은 광류의 지상 기준점으로부터 계산이 된다. 광류의 지상 기준점은 카메라 외부표정에 의해 계산 할 수 있으므로, 본 장에서는 각 영상 프레임에서 카메라 외부표정 계산방법을 구술한다. 카메라 외부표정 계산은 첫 번째 영상 프레임을 기준으로 설정하고, 두 번째 영상프레임의 외부표정은 공선조건을 기반으로 한 상호표정으로

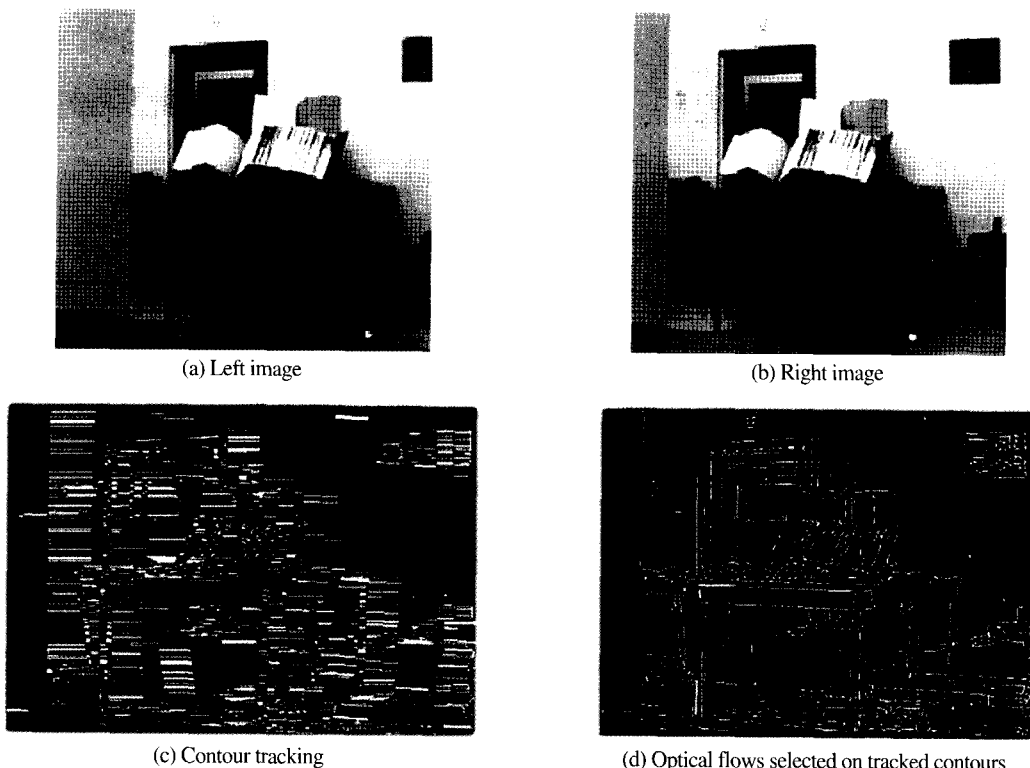


Fig. 5. Selecting the best matched optical flows.

구하며, 3번째 영상 프레임부터 마지막 영상 프레임은 광류의 지상 기준점을 대입한 절대표정으로 계산한다.

1) 공선조건을 기반으로한 상호표정

Fig. 6과 같이 영상 프레임은 카메라 초점위치  $o(X_L, Y_L, Z_L)$ 에서 카메라 좌표공간( $X' Y' Z'$ )의  $Z'$ 축으로 카메라 초점  $f$  거리에 위치하고, 직각좌표계(Cartesian coordinate system)공간에서  $P(X, Y, Z)$ 는 영상프레임의  $p(x, y)$ 에 투영된다.

$$\begin{pmatrix} x \\ y \\ f \end{pmatrix} = \lambda M \begin{pmatrix} X - X_L \\ Y - Y_L \\ Z - Z_L \end{pmatrix} \quad (3)$$

여기서  $\lambda$ 는 스케일이며,  $M$ 은 카메라 회전  $3 \times 3$ 행렬이다.

$$M = \begin{pmatrix} \cos \phi \cos x & \cos \phi \sin x + \sin \phi \sin \phi \cos x & \sin \phi \sin x - \cos \phi \sin \phi \cos x \\ -\cos \phi \sin x & \cos \phi \cos x - \sin \phi \sin \phi \sin x & \sin \phi \cos x + \cos \phi \sin \phi \sin x \\ \sin \phi & -\sin \phi \cos \phi & \cos \phi \cos \phi \end{pmatrix} \quad (4)$$

식(3)의 3번째 행으로 1, 2번째 행을 나누면 스케일  $\lambda$ 가 제거되고 일반적인 공선방정식이 유도된다.

$$F_x = x - f \frac{M_{11}(X - X_L) + M_{12}(Y - Y_L) + M_{13}(Z - Z_L)}{M_{31}(X - X_L) + M_{32}(Y - Y_L) + M_{33}(Z - Z_L)} \quad (5)$$

$$F_y = y - f \frac{M_{21}(X - X_L) + M_{22}(Y - Y_L) + M_{23}(Z - Z_L)}{M_{31}(X - X_L) + M_{32}(Y - Y_L) + M_{33}(Z - Z_L)}$$

첫 번째 영상 프레임을 Cartesian 좌표공간의 기준으로 설정하고, 두 번째 영상 프레임의 카메라 외

부표정은 식(5)에 Taylor 급수를 적용하여 식(6)와 같이 선형화 한 후에, 식(8)와 같이 전형적인 최소제곱법(Least-square)을 수 차례의 반복(Iteration)을 통하여 구할 수 있다.

$$F = \bar{F} + J^e \Delta^e + J^s \Delta^s + Error \quad (6)$$

여기서,  $\bar{F} = -F(\bar{\omega}, \bar{\phi}, \bar{x}, X_L^{manual}, \bar{Y}_L, \bar{Z}_L, x, y, f, \bar{X}, \bar{Y}, \bar{Z})$ ,  $\bar{F}_{4n \times 1} = [\dots \bar{F}_{x,t,i}, \bar{F}_{y,t,i}, \bar{F}_{x,t,i}, \bar{F}_{y,t,i}]$ , 외부표정 보정 값  $\Delta^e_{5 \times 1} = [\Delta\omega_t, \Delta\phi_t, \Delta x_t, \Delta Y_{L,t}, \Delta Z_{L,t}]$ , 광류의 지상 기준점 보정 값  $\Delta^s_{3 \times 1} = [\dots \Delta X_i, \Delta Y_i, \Delta Z_i, \Delta Z_{i+1}]$ .

$$J^e_{4n \times 5} = \begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \frac{\partial F_{x,t,i}}{\partial \omega_t} & \frac{\partial F_{x,t,i}}{\partial \phi_t} & \frac{\partial F_{x,t,i}}{\partial x_t} & \frac{\partial F_{x,t,i}}{\partial Y_L} & \frac{\partial F_{x,t,i}}{\partial Z_L} \\ \frac{\partial F_{y,t,i}}{\partial \omega_t} & \frac{\partial F_{y,t,i}}{\partial \phi_t} & \frac{\partial F_{y,t,i}}{\partial x_t} & \frac{\partial F_{y,t,i}}{\partial Y_L} & \frac{\partial F_{y,t,i}}{\partial Z_L} \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

$X_L^{manual}$ 수동입력으로서 본 논문에선 광류의 영상  $x$ 축 평균 이동량으로 설정하였다.  $t$  영상 프레임 번호,

$$J^s_{4n \times 3n} = \begin{pmatrix} \dots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots \\ \dots & \frac{\partial F_{x,t,i}}{\partial X_i} & \frac{\partial F_{x,t,i}}{\partial Y_i} & \frac{\partial F_{x,t,i}}{\partial Z_i} & 0 & 0 & 0 & \dots \\ \dots & \frac{\partial F_{y,t,i}}{\partial X_i} & \frac{\partial F_{y,t,i}}{\partial Y_i} & \frac{\partial F_{y,t,i}}{\partial Z_i} & 0 & 0 & 0 & \dots \\ \dots & 0 & 0 & 0 & \frac{\partial F_{x,t,i+1}}{\partial X_{i+1}} & \frac{\partial F_{x,t,i+1}}{\partial Y_{i+1}} & \frac{\partial F_{x,t,i+1}}{\partial Z_{i+1}} & \dots \\ \dots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots \end{pmatrix}$$

$i$ 는 광류 번호이다. 식(6)은 행렬형태로 표현하면 다음과 같다.

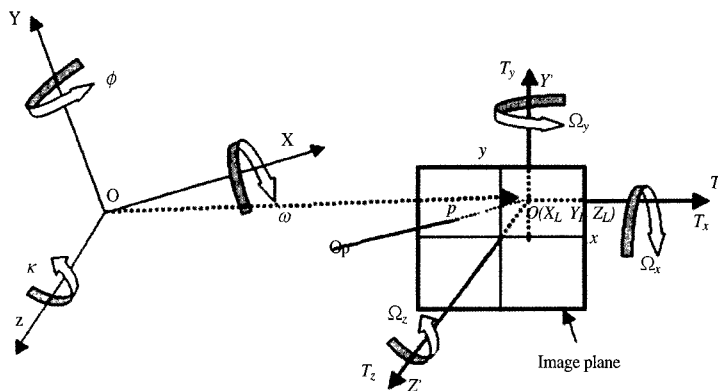


Fig. 6. Overview of perspective projection.

$$\begin{bmatrix} N_{11} & N_{12} \\ N_{12}^T & N_{22} \end{bmatrix} \begin{bmatrix} \Delta^e \\ \Delta^s \end{bmatrix} = \begin{bmatrix} n_1 \\ n_2 \end{bmatrix} \quad (7)$$

여기서,  $N_{11} = J^e T$ ,  $N_{12} = J^e T J^s$ ,  $N_{22} = J^s T J^s$ ,  $n_1 = J^e T \bar{F}$  그리고  $n_2 = J^s T F$  이다. 외부표정 보정  $\Delta^e$ 는 식(7)에서 광류 지상 기준점 보정값  $\Delta^s$ 를 제거하여 얻을 수 있다.

$$(N_{11} - N_{12} N_{22}^{-1} N_{12}^T) \Delta^e = n_1 - N_{12} N_{22}^{-1} n_2 \quad (8)$$

식(8)을 몇 번 반복하여, 카메라 외부표정  $[a_2 \phi_2 x_2 Y_{L2} Z_{L2}] = [\bar{a}_2 \bar{\phi}_2 \bar{x}_2 \bar{Y}_{L2} \bar{Z}_{L2}] - \Delta^e$ 을 구할 수 있다.

## 2) 광류의 지상 기준점을 대입한 절대표정

Fig. 7(a)와 같이 이전 두 영상 프레임의 카메라 외부표정을 이용하여 광류의 지상 기준점을 구할 수 있다. 이 광류의 지상 기준점을 공선방정식에 대입하여 카메라 외부표정을 3번째 영상 프레임부터 마지막 영상프레임까지 구할 수 있다. 여기서 필요로 하는 광류는 최소 3개 이상이 되어야 하며, 알고리즘을 만족하기 위해서는 최소 3개의 광류는 Fig. 7.b와 같이 반드시 3개의 연속된 영상

프레임에서 추적되어야 한다.

광류의 지상 기준점 계산은 다음과 같다.

$$\begin{aligned} X_i - X_{L,t-2} &= (Z_i - Z_{L,t-2}) \frac{\bar{U}_{t-2}}{\bar{W}_{t-2}} \\ Y_i - Y_{L,t-2} &= (Z_i - Z_{L,t-2}) \frac{\bar{V}_{t-2}}{\bar{W}_{t-2}} \\ X_i - X_{L,t-1} &= (Z_i - Z_{L,t-1}) \frac{\bar{U}_{t-1}}{\bar{W}_{t-1}} \\ Y_i - Y_{L,t-1} &= (Z_i - Z_{L,t-1}) \frac{\bar{V}_{t-1}}{\bar{W}_{t-1}} \end{aligned} \quad (9)$$

여기서,  $\bar{U} = M_{11}x_i + M_{21}y_i + M_{31}f$ ,  $\bar{V} = M_{12}x_i + M_{22}y_i + M_{32}f$ ,  $\bar{W} = M_{13}x_i + M_{23}y_i + M_{33}f$ 이다. 식(9)에서 광류의 지상기준점의 좌표는 최소 자승법으로 구할 수 있다.

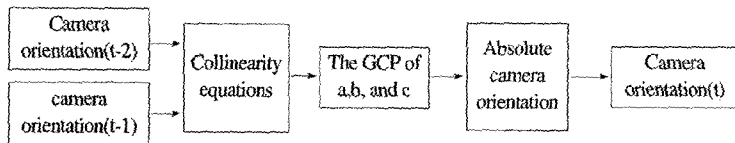
$$[X_i Y_i Z_i] = (A^T A)^{-1} A^T B \quad (10)$$

여기서,

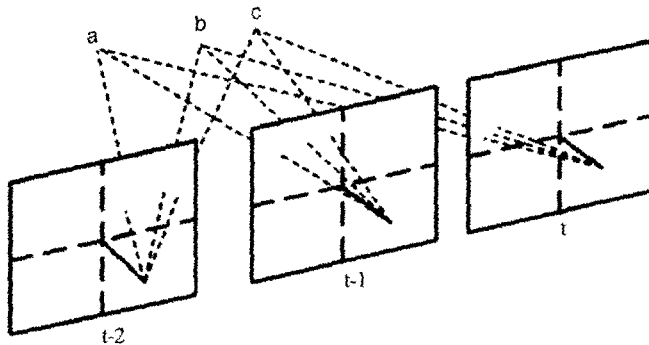
$$A = \begin{bmatrix} 1 & 0 & -\bar{U}_{t-2}/\bar{W}_{t-2} \\ 0 & 1 & -\bar{V}_{t-2}/\bar{W}_{t-2} \\ 1 & 0 & -\bar{U}_{t-1}/\bar{W}_{t-1} \\ 0 & 1 & -\bar{V}_{t-1}/\bar{W}_{t-1} \end{bmatrix}, \quad B = \begin{bmatrix} X_{L,t-2} - Z_{L,t-2} \frac{\bar{U}_{t-2}}{\bar{W}_{t-2}} \\ Y_{L,t-2} - Z_{L,t-2} \frac{\bar{V}_{t-2}}{\bar{W}_{t-2}} \\ X_{L,t-1} - Z_{L,t-1} \frac{\bar{U}_{t-1}}{\bar{W}_{t-1}} \\ Y_{L,t-1} - Z_{L,t-1} \frac{\bar{V}_{t-1}}{\bar{W}_{t-1}} \end{bmatrix}$$

이다. 계산된 광류의 지상 기준점을 절대표정 식에 대입하여 3번째 영상 프레임부터 마지막 영상 프레임까지의 카메라 외부표정을 구할 수 있다.

$$\Delta^e = (J^e T J^e)^{-1} J^e T \bar{F} \quad (11)$$



(a)



(b)

Fig. 7. Camera translation through tracked 3 optical flows during 3 image frames.

식(11)를 몇 번 반복하여, 카메라 외부표정  $[\omega_i \phi_i x_i X_{L,i} Y_{L,i} Z_{L,i}] = [\bar{\omega}_i \bar{\phi}_i \bar{x}_i \bar{X}_{L,i} \bar{Y}_{L,i} \bar{Z}_{L,i}] - \Delta^e$ 를 구할 수 있다.

#### 4. 영상 모자이크를 위한 다중 기선

앞에서 언급한 것과 같이 2차원 영상 모자이크는 작은 여러 영상을 하나의 기선에 투영하여 넓은 대형영상을 제작하는 것이다. 2차원 영상 모자이크는 카메라 이동이 도심지역의 도로를 따라 움직일 때는 하나의 기선에 영상 모자이크를 생성할 수 없다. 본 연구에서는 Fig. 1(b)와 같은 경우에 있어서도 영상 모자이크가 가능한 기법을 제안하였다. 제안한 방법은 각 영상 프레임마다 자신의 기선을 가지는 다중 기선을 구하여 그 영상을

투영하는 것이다. 기선 결정은 계산된 광류의 공간좌표를 1차 방정식으로 결정된다. 본 연구에서 카메라의 Y축은 도심지역의 건물의 높이방향과 유사하다고 가정하며, 카메라의 X축 이동이 카메라의 주 움직임 된다. 따라서 기선을 Fig. 8(a)와 같이 XZ평면에서 계산할 수 있다. Fig. 8(b)와 같은 상황에서 기선은 굵은 점선이 된다. 각 기선에 해당되는 영상 프레임을 투영하면, 3차원 공간에서 영상 모자이크는 Fig. 8(c)와 같은 결과가 얻어지게 된다.

#### 1) 영상 모자이크 생성이 어려운 기선보정

Fig. 8b의 예는 기선간들의 만나는 점이 이상적인 경우일 때이다. 일반적으로 Fig. 9와 같이 기선간에 평행하여 만나지 않거나, 기선간에 잘못된 지점에서 만나는 경우가 빈번히 발생한다.

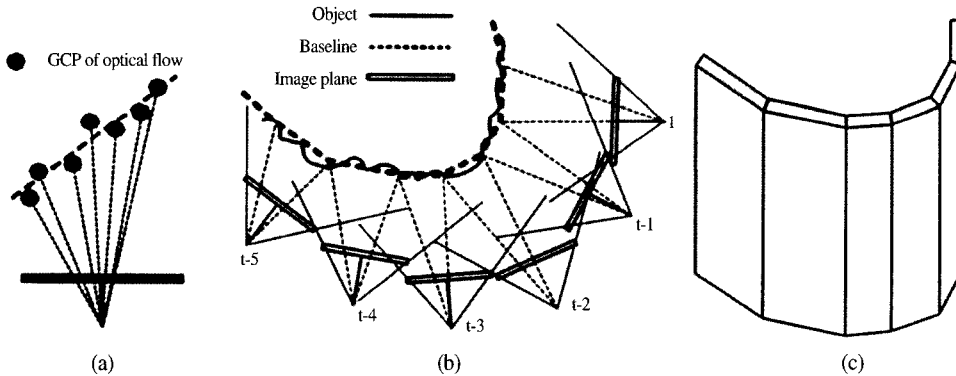


Fig. 8. Independent baselines by using camera pose and average depth of feature points per image frame.

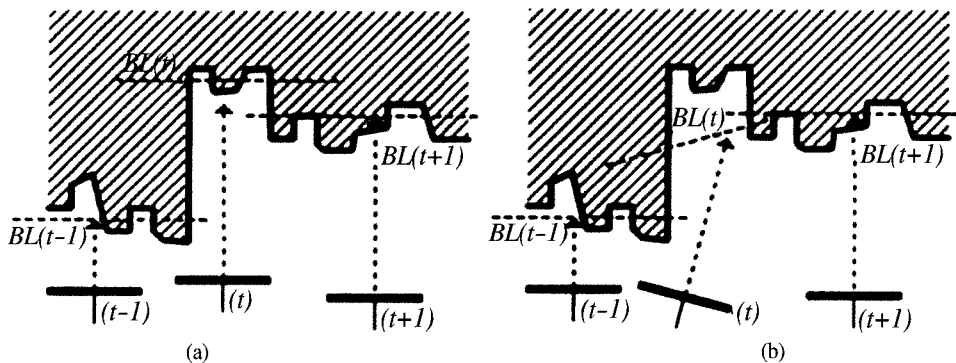


Fig. 9. Parallel baselines and fault intersection among baseline(BL).



이 단점을 보정하기 위하여, 다음 3단계 과정을 거쳐 기선을 보정이 된다.

Step 1: 모든 영상 프레임에 대한 광류의 지상 기준점을 대표하는 1차 방정식  $Z=a_tX+b_t$ 을 구한다.

Step 2: 짝수 영상 프레임에 해당하는 1차 방정식  $Z=a_tX+b_t$ 이 이전 홀수 영상 프레임의 영상 중심화소가  $Z=a_{t-1}X+b_{t-1}$ 에 투영되는 점 A부터  $T_{X,t-1}/2$ 가  $Z=a_{t-1}X+b_{t-1}$ 에 투영되는 점까지 지나가는지를 검사한다. 이 영역을 지나가지 않을 경우엔, Fig. 10과 같이  $B=A+2(A-C)/5$  을 기선이 만나는 점으로 설정한다.

Step 3: Step 2와 같은 방법으로, 짝수 영상 프레임에 해당하는 1차 방정식  $Z=a_tX+b_t$ 이 이후 홀수 영상 프레임의 영상 중심화소가  $Z=a_{t+1}X+b_{t+1}$ 에 투영되는 점 F부터  $-T_{X,t+1}/2$ 가  $Z=a_{t+1}X+b_{t+1}$ 에

투영되는 점까지 지나가는지를 검사한다. 이 영역을 지나가지 않을 경우엔, Fig. 10과 같이  $E=F+2(F-D)/5$  을 기선이 만나는 점으로 설정한다. 새로운 짝수 기선은 Fig. 10과 같이  $BE$ 를 지나가는 선이 된다. 여기서 A, C계산은 다음과 같다.

$$A_X = \frac{\{X_{L,t-1} + (b_{t-1} - Z_{L,t-1})M_{31,t-1}/M_{33,t-1}\}}{\{1 - a_{t-1}M_{31,t-1}/M_{33,t-1}\}}$$

$$A_Z = a_{t-1}A_X + b_{t-1}$$

$$C_X = \frac{\{X_{L,t} + (b_{t-1} - Z_{L,t})M_{31,t}/M_{33,t}\}}{\{1 - a_{t-1}M_{31,t}/M_{33,t}\}}$$

$$C_Z = a_{t-1}C_X + b_{t-1}$$
(12)

D, F는 식(12)와 같이 계산하면 된다. 일반 기선과 보정된 기선을 연결하면, Fig. 1b와 같은 도심지에서의 기선은 Fig. 11과 같은 결과를 얻을 수 있을 것이다.

### 2) 기선에 대한 영상프레임의 투영

계산된 기선에 대한 영상프레임의 투영은 식 (5)을 이용하여 구할 수 있다. Fig. 12에서 a, b, c, d의 X, Z는 식(12)에 의해 계산되며, a, b, c, d의 Y

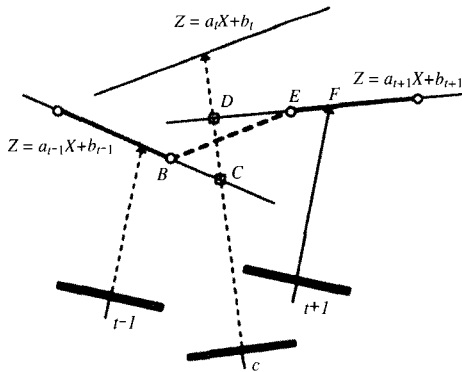


Fig. 10. New baseline in case of fault intersection among baselines (BL).

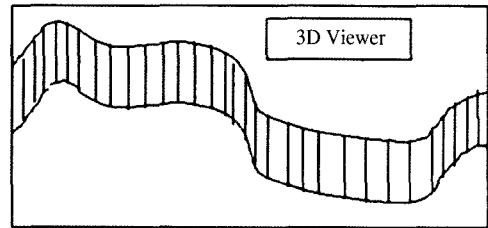


Fig. 11. Conceptual result by the proposed method.

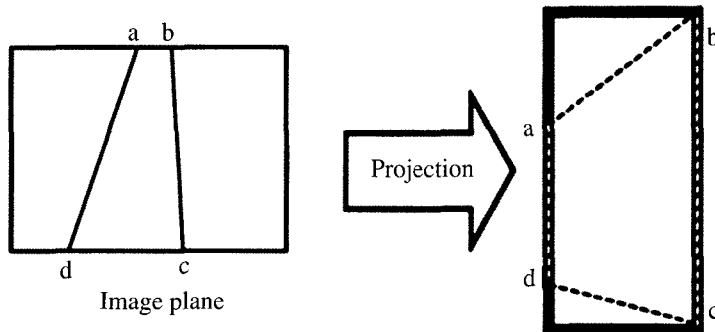


Fig. 12. Projection relationship between an image plane and an base-line.

는 다음 식으로 계산된다.

$$Y = \frac{(fM_{21}-yM_{31})(X-X_L) + (yM_{32}-fM_{22})Y_L + (fM_{23}-yM_{33})(Z-Z_L)}{yM_{32}-fM_{22}} \quad (13)$$

여기서  $y$ 는 영상의 상,하단부분을 의미한다. 영상 좌표에서  $a, b, c, d$ 의  $x$ 는 다음 식으로 계산된다.

$$x = f \frac{M_{11}(X-X_L) + M_{12}(Y-Y_L) + M_{13}(Z-Z_L)}{M_{31}(X-X_L) + M_{32}(Y-Y_L) + M_{33}(Z-Z_L)} \quad (14)$$

### 5. 기선간 만나는 경계를 인식하지 못하는 화소

카메라 회전에 의해서 얻어진 반구형, 파노라믹 모자이크 경우, seamless 영상 모자이크 기법이 중요

하지 않으므로 영상 블렌딩(blending) 알고리즘인 가중치, 저 주파수 블렌딩 기술로 miss-alignment 문제를 해결 할 수 있으나, 카메라 이동에 의해서 영상 모자이크는 두 영상간의 화소차이가 심하게 나타므로, 모자이크된 영상들의 경계 부분에서 사람의 눈으로 보았을 때 연결되었다는 것을 인식 못하게 하는 seamless 영상 모자이크가 매우 중요하다. 이러한 현상이 발생하는 이유는 카메라의 이동 (스테레오 스코픽)이 있기 때문이다. 예를 들면, Fig. 13(a)와 같이 두 기선이 만나는 점에서 기선과 영상에 투영된 물체와 지리적인 불일치일 경우, 왼쪽 영상에 물체 A가 투영되며, 오른쪽 영상에 물체 B가 투영되어 모자이크된 영상들의 경계부분에서 영상 불일치가 발생한다. 이 경우엔 물체 C부분을 경계로 할 때 좋은 결과를 얻을 수 있다. 즉 비슷한 화소들을 찾아 연결한 곡선을 이

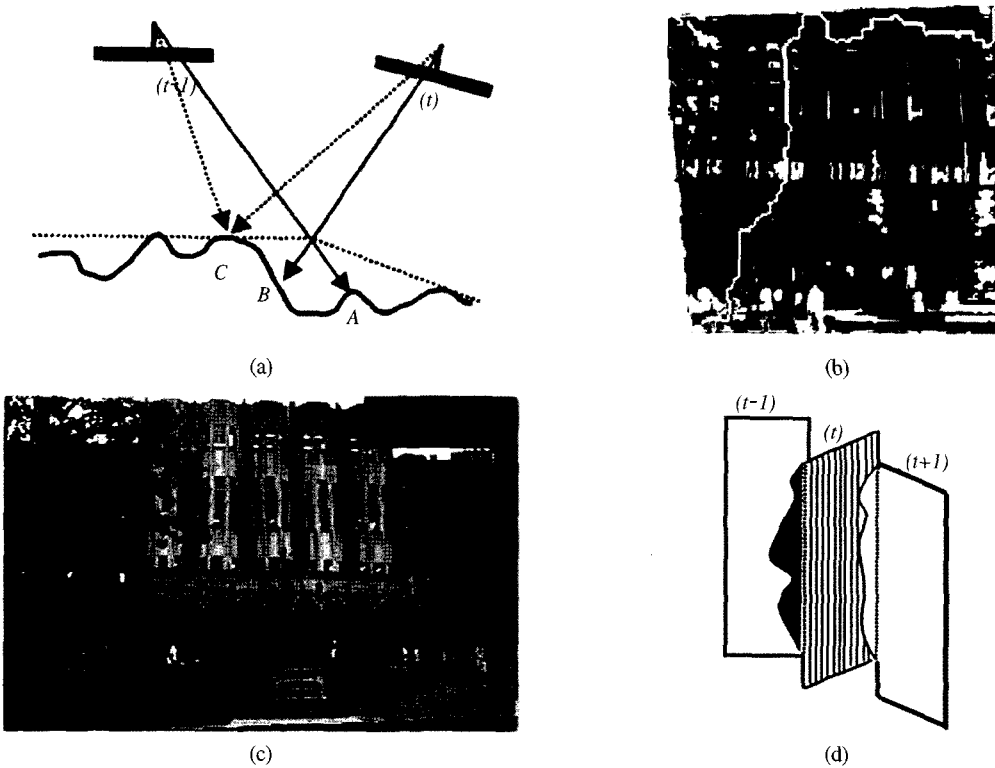


Fig. 13. Seamless image mosaics using seam-line detected by the dynamic programming, (a) the reason of miss-alignment overlapping area between two images, (b) seam-line detected by the dynamic programming on correlation map in overlapping area, (c) the seamless mosaics using the seam-line, and (d) concept of seamless video mosaics in 3D space.

용하여 모자이크를 하는 것이다. 이 곡선은 하나의 경로로 볼 수 있으며, 최소경로를 찾는 방법인 동적계획법을 적용 할 수 있다. 현재 영상 프레임을 이전 영상 프레임의 기선에 투영한 것과 이전 영상과 상관정합을 한 후 시작점에서 도착점에 이르는 최소경로를 동적 계획법으로 찾는다. 두 겹친 영상을 상관정합을 비용지도로 만든 후에 동적 계획법을 적용하면 Fig. 13(b)와 같이 굵은 곡선을 얻을 수 있다. 이 곡선을 기반으로 영상 모자이크를 생성하면 Fig. 13(c)와 같이 생성된다. 이러한 개념을 제안한 방법에 대입하면 Fig. 13(d)와 같은 결과를 얻을 수 있다.

### 6. 3차원 다중 기선을 이용한 영상 모자이크 구현

본 논문에서 제안한 3차원 다중 기선을 이용한 영상 모자이크 기법을 실제 영상 프레임에 적용하였다. 제안한 알고리즘은 *Visual C++ 6.0*으로 구현하였으며, 알고리즘 결과물의 디스플레이는 3D 디스플레이 라이브러리인 *OpenGL*을 사용하였다. Fig. 14의 영상 프레임은 디지털 카메라를 회전하면서 이동시켜 촬영한 영상이며, 디지털 카메라의 포커스  $f$ 는 1.0으로 가정하였고, 영상 크기는 1280 × 960이다. Fig. 15(a)는 연속된 영상프레임의 카메라 외부표정과 광류의 공간좌표를 표현 나타내고 있다. Fig. 15(b)는 계산된 기선에 영상 프레임을

나타낸 것이다. Fig. 15(c)는 계산된 기선을 보여주고 있으며 15(d)는 기선을 *XZ*평면에서 보여주고 있다. Fig. 16(a)는 제안한 알고리즘을 구현한 소프트웨어이며, 16(b)는 건물의 왼쪽 면을 보여주고 있으며, Fig. 16(c)는 건물의 정면, Fig. 16(d)는 3D 공간에서 영상 모자이크된 결과물을 위쪽 방향에서 아래 방향으로 표현한 결과이다.

## 7. 결론

본 논문에서는 비데오 카메라가 도심지역의 도로를 따라서 이동하여 취득된 영상에 대하여 3차원 공간에서 영상 모자이크를 생성 할 수 있는 새로운 기법을 제안하였다. 본 연구에서 제안된 알고리즘은 각 영상 프레임마다 독립적인 기선을 가지도록 하여 3차원 공간에서 영상 모자이크 생성하는 방식으로서, 이는 계층적 방법을 이용한 영상 프레임 간의 특징점 추출 및 이들을 정합한 광류계산, 광류를 이용한 카메라 외부 표정 계산 및 각 영상프레임의 광류의 지상 기준점들을 1차 방정식으로 표현에 의한 다중 기선계산의 과정을 통해 구현되고 있다. 제안된 기법을 실제 영상 프레임에 적용하여 실험한 결과, 이동요소와 회전요소를 가진 연속적인 영상 프레임에 대해서도 효과적으로 3차원공간에서 영상 모자이크를 구현하고 있음을 알 수 있었다. 향후, 본 연구에서 제안된 기법을 차량을 이용한 모바일 맵핑 시스템(vehicle-

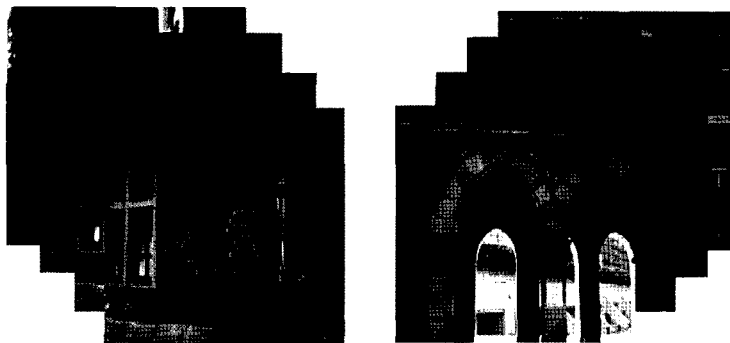


Fig. 14. Real sequential image frames taken from a moving and rotating digital camera.

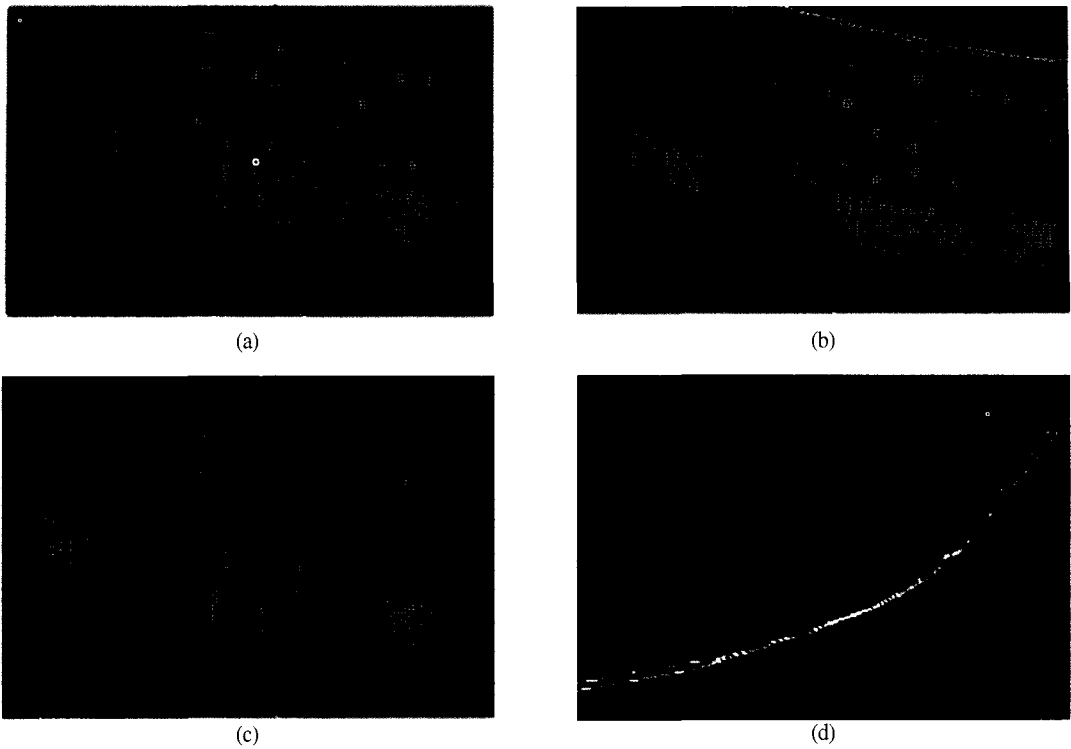


Fig. 15. 3D multi-baselines, red sphere is GCP of optical flows and green square is image plane, (a) relationship of a moving camera and GCP, (b) Image planes projected on itself dependent multi-baseline, (c) multi-baselines in 3D space, (d) multi-baseline in XZ plan coordinate.

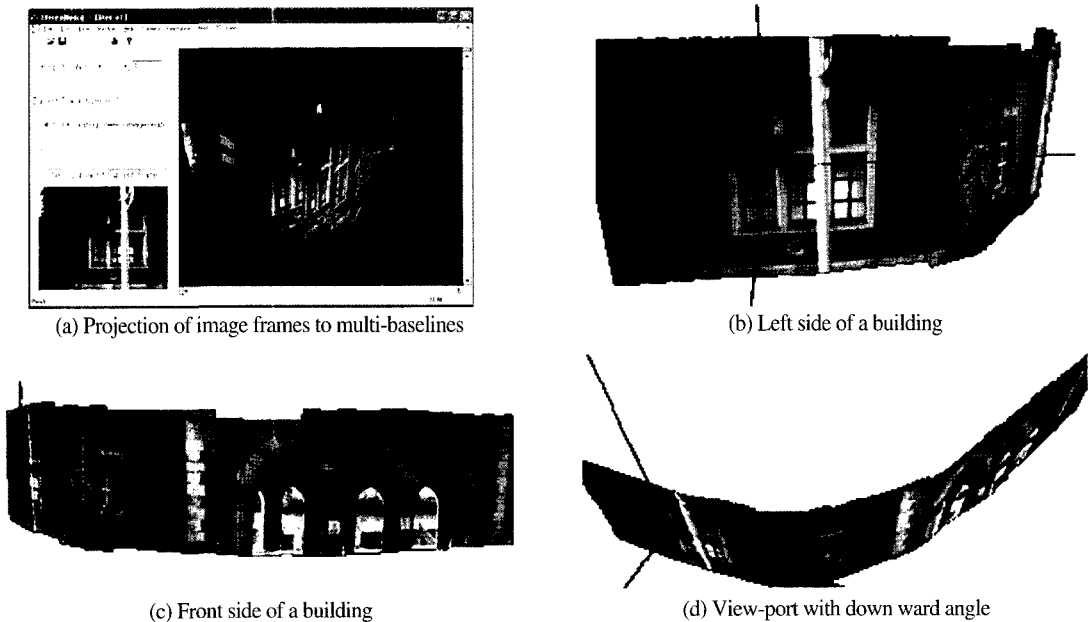


Fig. 16. Projection of image frames to multi-baselines.

borne mobile mapping system)과 같이 이동요소와 회전요소를 동시에 가지는 시스템의 CCD 카메라 영상에 적용한다면, 기존의 레이저 스캐너를 대체할 수 있는 저비용, 고효율의 효과적인 3차원 공간데이터 취득 시스템 제작이 가능할 것으로 사료되며, 또한 제안한 방법에 의해 생성된 영상 모자이크는 영상의 연속성이 부여되어 있으므로, 기존의 파노라믹 과 반구형 영상 모자이크를 이용한 가상현실 구현에서 문제점으로 발생되고 있는 영상과 영상사이에 데이터의 부족으로 인한 영상단절 현상을 효율적으로 극복하고, 보다 현실적인 가상현실 구현과 공간정보의 제공이 가능할 수 있을 것으로 판단된다.

## 참고문헌

- CHON, J. and H. KIM, 2001. The computation of the real-time CCD camera motion using correlation and system dynamic model, Proc. of 2001 *IEEE International Conference on M2VIP*, Hongkong, 27(8): 168-172.
- Chou, J. S., and J. Qian, Z. Wu, H. Schramm, 1997. Automatic mosaic and display from a sequence of peripheral angiographic images, *Proc. of SPIE, Medical Imaging, California*, 3034: 1077-1087.
- Coorg, S. and S. Teller, 1999. Extracting textured vertical facades from controlled close-range imagery, *Proc. of 1999 IEEE Conference on Computer Vision and Pattern Recognition*, Colorado, pp. 625-632.
- Coorg, S. and S. Teller, 2000. Spherical mosaics with quarterings and dense correlation, *In IJCV*, 37(3): 259-273.
- Sawhney, H. and S. Ayer, 1997. Compact representation of video through dominant and multiple motion estimation, *IEEE Trans. PAMI*, 18: 814-830.
- Han, J. and J. Park, 2000. Contour Matching Using Epipolar Geometry, *IEEE Trans. PAMI*, 22(4): 358-370.
- Shum, H. and R. Szeliski. 2000. Systems and experiment paper construction of panoramic image mosaics with global and local alignment, *In IJCV*, 36(2): 101-130.
- Smith, S. M. and J. M. Brady, 1997. SUSAN-a new approach to low level image processing, *In IJCV*, 23(1): 45-78.
- Szeliski, R., 1996. Video mosaic for virtual environment, *Comput. Graph. Applicat.*, 16(3): 22-30.
- USGS Hurricane Mitch Program Projects, <http://mitchnts1.cr.usgs.gov/projects/aerial.html>
- Zhu, Z., A. R. Hanson, H. Schultz, and E. M. Riseman, 2001. Error Characteristics of Parallel-Perspective Stereo Mosaics, *Proc. of IEEE Workshop on Video Registration (with ICCV01)*, Vancouver, Canada, 13(7).