

# 가변 분해능을 가진 2차원 히스토그램을 이용한 강건한 광류검출

## Robust Optical Flow Detection Using 2D Histogram with Variable Resolution

전재춘<sup>1)</sup>

CHON, Jaechoon

### Abstract

The proposed algorithm is to achieve the robust optical flow detection which is applicable for the case that the outlier rate is over 80%. If the outlier rate of optical flows is over 30%, the discrimination between the inliers and outlier with the conventional algorithm is very difficult. The proposed algorithm is to overcome such difficulty with three steps of grouping algorithm; 1) constructing the 2 D histogram with two axes of the lengths and the directions of optical flows. 2) sorting the number of optical flows in each bin of the two-dimensional histogram in the descending order and removing some bins with lower number of optical flows than threshold. 3) increasing the resolution of the two-dimensional histogram if the number of optical flows in a specific bin is over 20% and decreasing the resolution if the number of optical flows is less than 10%. Such processing is repeated until the number of optical flows falls into the range of 10%-20% in all the bins. The proposed algorithm works well on the different kinds of images with many of wrong optical flows. Experimental results are included.

Keywords : Optical flow, Outlier, Inlier, Moving objects, Tracking

### 요 지

두 영상에서 정확한 광류를 얻고자 할 경우, 잘못 정합된 광류의 수가 전체의 80%이상일 경우에도, 정합이 잘된 광류만을 빠르고 안정적으로 추출하는 방법을 제안하였다. 제안한 방법은 다음과 같은 단계로 구성된 그룹화 알고리즘이다. 1) 광류의 방향과 길이성분으로 구성된 2차원 히스토그램을 구성한 후, 2) 2차원 히스토그램에서 광류수가 높은 그룹들부터 내림차 순으로 정렬한 후 각 그룹에서 광류수가 적은 그룹은 제거한다. 또, 3) 선택된 그룹들의 광류수가 20% 이상일때는 히스토그램의 분해율을 증가 시키고, 10% 이하일 경우에는 히스토그램의 분해율을 감소시켜서 위 과정을 반복한다. 선택된 그룹들의 광류수가 10%에서 20%사이에 올 때까지 이 반복을 계속한다. 제안한 알고리즘을 잘못된 광류가 많이 포함된 다양한 종류의 영상에 대해서 적용한 결과 강건하게 정합이 되어 잘못된 광류를 제거할 수 있었다. 이 논문은 이에 대한 실험결과를 포함하였다.

핵심용어 : 광류, outlier, inlier, 움직임 물체, 추적

## 1. 서 론

항공영상이나 동영상에서의 3차원 데이터 취득 혹은 영상 모자이크를 위해서는 화소 간의 동일한 위치 점들을 추적하여 광류(optical flow)를 추출하는 일이 필요하다. 이 광류 추출방법에는 Gradient based(Kearney et. al., 1987), Frequency based(Adelson and Bergen, 1985), 특

징점 기반 방법(Smith and Brady, 1995) 등이 있는데, Gradient based 방식은 Horn and Schunck 알고리즘에서와 같이 시간적으로 인접한 두 영상에서 최대의 영상 변화방향을 이용하는 방법이다(Horn and Schunck, 1981). 또, Frequency based는 Gabor filters(Adelson and Bergen, 1985)와 같이 속도에 대한 band-pass filter를 적용한 방법이다.

1) 정회원 · 동경대학 토목공학과 박사과정(E-mail:jjc7151@trip.t.u-tokyo.ac.jp)

두 방법을 이용한 광류계산에서는 빠른 계산이 어려울 뿐만 아니라 영상이동이 1픽셀 이내 일 경우에만 적용 가능하다는 단점이 있다.

광류를 이용하여 카메라 촬영방향이나 모션을 구하는 응용들에서는 영상의 모든 픽셀에서 광류가 필요하지 않고, 적은 수의 광류라고 하더라도, 정확한 광류를 이용하면 보다 더 정확한 카메라 모션을 계산할 수 있게된다. 특징점들을 기반으로 방법은 이와 같은 응용에 적합한 알고리즘인데, 화소의 움직임 파악이 용이한 특징점 화소들만을 대상으로 추출한 광류들을 이용한다. 이 때의 특징점으로는 모서리점(corner point)을 많이 이용하는데, 대표적인 모서리점 추출방법으로는 Plessey, Kitchen/Rosenfeld(Kitchen and Rosenfeld, 1982), Curvature Scale Space(Mokhtarian and Riku, 1998), SUSAN[8] 방법등이 있다. 이 중에서 SUSAN 방법은 공간영역에서 강도(intensity) 차이를 가우시안(Gaussian) operation을 적용하여 모서리점을 구하는 효과적인 방법이다(Smith and Brady, 1997). 특징점을 기반으로한 광류 추출 알고리즘들 중 빠른 정합방법으로는 상관정합인 Sum of Squared Difference(SSD)를 사용한다. 이 방법은 영상회전, 줌(zoom), 영상변짐(blur)현상 등이 발생할 때, 잘못 정합된 광류가 많아질 때, 많이 사용되는데 이 때, 잘못된 광류를 제거하는 방법이 중요하다. 잘못 정합된 광류를 제거하는 방법으로는 가중치 최소제곱법인 M-estimators(Haralick et al, 1998)와 Least Median of Squares(LMedS)(Rousseeuw, 1984)가 일반적으로 이용되고 있다. M-estimators는 일차적으로 카메라 모델에 추출된 광류를 대입하여 모델을 구하고, 이 모델에 광류의 시작점을 입력하여 추정된 위치와 실제 광류의 끝점과의 거리에 따라 가중치를 부여하여 광류를 다시 추출하는 방법이다. 그런데, 이 방법에서는 잘못 정합된 광류가 30%이상일 때는 계산이 불가능한 단점이 있다. 보다 강건한 방법으로 LMedS가 있는데, M-estimators 처럼 일차적으로 카메라 모델에 추출된 광류를 대입하여 모델을 구한후, 이 모델에 광류의 시작점을 입력하여 추정된 위치와 실제 광류의 끝점과의 거리가 어느 이상이 되면 광류유류로 간주하여 그 광류를 제거하는 방식이다. 이 알고리즘도 잘못 정합된 광류가 35%이상 일때는 적용이 어렵다는 문제점이 있다. 즉, 영상정보가 충분치 않은 영상에서는 잘못 정합된 광류가 더 많이 나타날 수 있다. 예를 들면, 회전과 줌이 클 경우나, 유사한 텍스처를 많이 포함한 영상의 경우, 영상변짐이 존재하는 경우등에서는 잘못 정합된 광류가 60% 이상 나타

날 수 있다. 이 경우에는 기존 방법만으로 문제를 해결하기 어렵다.

본 논문에서는 잘못 정합된 광류가 80%이상일 경우에도 잘못된 광류를 강건하게 제거 해낼수 있는 알고리즘을 제안했다. 영상에서 광류를 계산해 보면, 잘못된 광류들 끼리는 지역적으로 일관성을 유지하지만, 잘못 정합된 광류는 일관성이 없거나, 일관성을 유지하더라도 그 수가 많지 않는 점에 착안하여, 유사한 벡터를 그룹화하고 벡터수가 적은 그룹은 제거하는 방법을 사용하였다. 제안한 알고리즘은 다음과 같이 3단계로 구성되어있다. 즉, 1)광류의 방향과 길이성분으로 구성된 2차원 히스토그램을 구하고, 2)2차원 히스토그램을 광류수가 높은 그룹 순으로 나열한 후 각 그룹에서 광류수가 적은 그룹은 제거한다. 3)선택된 그룹들의 광류수가 20% 이상일때는 히스토그램의 분해수를 증가 시키고, 10%이하일 경우에는 히스토그램의 분해수를 감소 시켜 두 그룹을 합한다. 이 과정을 2번부터의 반복하여 선택된 그룹들의 광류수가 10%에서 20%사이일 때 정지한다.

제안한 알고리즘에 의해 선택된 광류의 정확도를 평가하기 위해 화소와 카메라 간의 관계를 메트릭스로 표현한 affine함수를 이용하였다. 즉, affine함수를 계산하여 카메라 움직임에 의해 나타난 배경 광류를 구한후, affine에 의해 추정된 위치와 다르게 움직이는 그룹들에서 광류가 일정 수 이상인 그룹을 선택하여 움직이는 물체로 인식하는데 응용하였다. 제안한 알고리즘은 비디오 동영상에서 각 프레임의 카메라 오리엔테이션 자동으로 계산에 응용될수 있다(장호식 외2인, 2003; 이진덕, 최용진, 2001).

본 논문의 구성은 2장에서 SUSAN 알고리즘과 상관정합 방법을 소개 하였고, 3장에서는 2차원 히스토그램을 이용한 광류의 그룹화 방법을 기술하며, 4,5장은 제안한 알고리즘을 적용한 실험과 결론으로 구성되어 있다.

## 2. 특징점 추출과 상관정합

### 2.1 SUSAN에 의한 특징점 추출

본 연구에서는 특징점 추출을 위해서 모서리점 추출에 성능이 좋은 SUSAN(Smallest Univalve Segment Assimilating Nucleus) 알고리즘을 이용하였다(Smith and Brady, 1997). SUSAN에서 사용한 경계 강도 계산은 원형 마스크 중심(그림 1에서 +) Univalve Segment Assimilating (USAN)의 영상 값의 원형 마스크 영역 내에 있는 각 영상

값과 차이를 합산한 결과 값을 가우시안(Gaussian) 함수를 적용하여 나타낸 것이다.

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

여기서, I는 화소(gray) 값이고 T는 화소 값의 차에 대한 임계값(threshold) 값이다. 식 (1)에서 T가 크면 화소 그래프 변화가 큰 곳만 특징 점으로 추출되고, T가 작으면 화소 변화가 작은 곳에서도 특징점을 추출한다. 특징점을 선택하는 기준은 G 값에 달려있다. 일정한 수의 특징 점들을 구하기 위해서는 T 값을 고정 시켜 지수함수 부분을 계산 후 G 값을 변화시키면 된다.

$$S = \begin{cases} S'; & \text{if } \sum e^{-\frac{|I(x,y) - I(x+dx, y+dy)|^6}{T}} < G \\ 0; & \text{Otherwise} \end{cases} \quad (2)$$

본 논문에서는 원형 마스크 지름의 크기를 7 화소, T는 30, G는 가변 하도록 설정 하였다. 그림 1(a)는 3가지의 원형 마스크를 보여주고 있다. 3가지 형태에서 모서리 점을 뜻하는 a형의 Smallest Univalence Segment Assimilating Nucleus(SUSAN) 값이 가장 작다. 그림 1(b)는 원형 마스크의 중심(USAN)이 흰색인 경우이다. 경계 강도에서 특징점을 추출하기 위하여 지역최대(Local Max)를 사용했으며, 추출된 특징점의 수에 따라 G값이 변하게 된다.

식 (1)에서는 지수함수를 영상의 각 픽셀에 적용하기 때문에 많은 계산 시간을 필요로 하게 되므로, 빠른 계산을 위해서 일반적으로 SUSAN상수의 T를 고정하며 지수함수 부분을 look-up테이블화 한다. 하나의 밴드를 가지는 영상에 대해서는 이 지수함수 부분은 256개의 배열인 look-up 테이블이 된다. 그림 2는 SUSAN 알고리즘을 적용 했을 때 결과를 보여주고 있다.

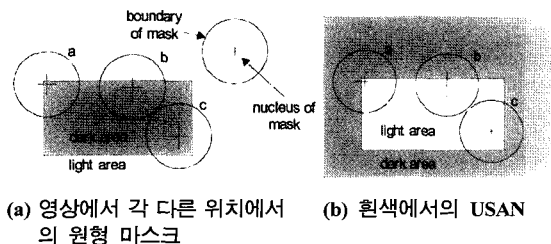


그림 1. SUSAN 원리

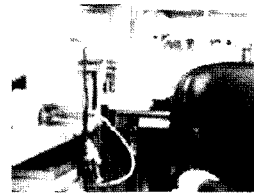
## 2.2 특징점 정합

특징점 정합방법으로는 주로 상관정합을 이용하는데, 빠른정합을 위하여 (3)과 같이 Sum of Absolute Differences (SAD)와 Sum of Squared Differences(SSD)를 사용한다.

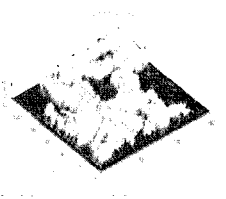
$$SAD = \sum_{i=-N/2}^{N/2} \sum_{j=-N/2}^{N/2} |I_1(x+i, y+j) - I_2(x+i+dx, y+j+dy)|$$

$$SSD = \sum_{i=-N/2}^{N/2} \sum_{j=-N/2}^{N/2} (I_1(x+i, y+j) - I_2(x+i+dx, y+j+dy))^2 \quad (3)$$

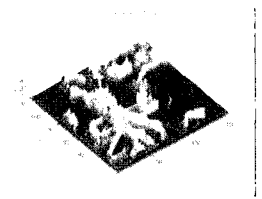
여기서,  $I_1, I_2$ 는 현재와 이전 영상 프레임, N은 마스크 크기이다. 상관정합의 그 외 방법으로는 Cross Correlation (CC), Normalized Cross Correlation(NCC)과 Zero-mean Normalized Cross Correlation(ZNCC)방법 등이 있으나, 계산 소요 시간이 많이 소요되기 때문에, 본 논문에서는 SSD 방법을 사용하였다.



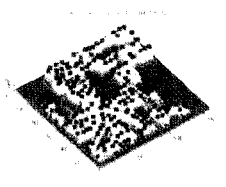
(a) 원 영상



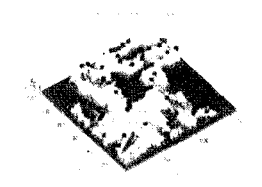
(b) T가 20일때 SUSAN 경계 강도



(c) T가 70일때 SUSAN 경계 강도



(d) T, G가 각각 30, 30일때 174 개의 특징점



(e) T, G가 각각 30, 45일때 52 개의 특징점

그림 2. SUSAN의 상수변화에 의한 경계강도와 특징점

### 2.3 카메라 동작과 화소 간의 관계를 위한 affine 함수

정확한 optical flow 여부를 결정하기 위해서, 본 논문에서는 공간상의 평면이 회전, 전이, 크기 변화와 이들의 조합으로 이루어진 선형기반 변환인 Affine 함수를 사용하였다.

$$\begin{aligned} x_2 &= a_1x_1 + a_2y_1 + a_3 \\ y_2 &= a_4x_1 + a_5y_1 + a_6 \end{aligned} \quad (4)$$

여기서,  $(x_1, y_1)$ 과  $(x_2, y_2)$ 는 기준영상과 다음 영상프레임의 영상좌표를 뜻한다. Affine식은 미지수가 6개( $a_1, a_2, a_3, a_4, a_5, a_6$ )이므로, 두 영상간의 동일점이 최소 3개가 있어야 이 미지수를 구할 수 있다. 식 (9)를 행렬로 표현하면 다음과 같다.

$$B = Ap \quad (5)$$

여기서,

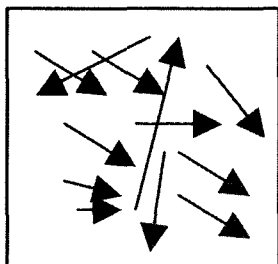
$$B = \begin{bmatrix} x_{2,i} \\ x_{2,j} \\ \vdots \\ x_{2,i+n} \\ x_{2,j+n} \end{bmatrix}, \quad A = \begin{bmatrix} 1 & x_{1,i} & y_{1,i} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & x_{1,j} & y_{1,j} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{1,i} & y_{1,i} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & x_{1,j} & y_{1,j} \end{bmatrix}, \quad p = \begin{bmatrix} a_3 \\ a_2 \\ a_1 \\ a_6 \\ a_5 \\ a_4 \end{bmatrix}$$

이며,  $i$ 와  $j$ 는 각각 기준영상과 다음 프레임의 영상에서 추출되어 정합된 특징점의 번호이다. 식 (5)는 최소제곱법을 사용하여 6개의 미지수를 구할 수 있다.

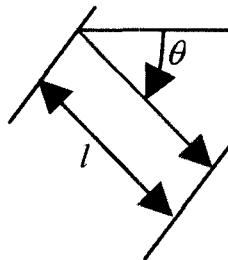
$$p = (A^{-1}A)A^{-1}B \quad (6)$$

### 3. 유사한 광류의 그룹화

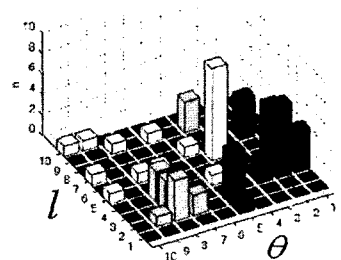
정합된 특징점들에서는 그림 3(a)와 같이 벡터성분을 갖는 광류를 얻을 수 있는데, 이 광류는 그림 3(b)와 같이 방향  $\theta$ 와 길이  $l$ 로 분해할 수 있다. 이와 같은 광류들에서 두 성분을 가로와 세로축으로한 2차원 히스토그램에 광류의 수를 누적하여 3차원 공간에서 도시하면 그림 3(c)와 같은 결과를 얻을 수 있다. 이 2차원 히스토그램에서 정합이 잘된 광류를 추출하는 알고리즘은 그림 6의 순서도에서와 같이 다음 3단계로 구성되어 있다. 1) 광류의 방향과 길이 성분으로 구성된 2차원 히스토그램을 구한다. 2) 2차원 히스토그램에서 광류수가 높은 그룹 순으로 나열한 후 각 그룹에서 광류수가 적은 그룹은 제거한다. 3) 선택된 그룹들의 광류수가 20% 이상일 때는 히스토그램의 분해수를 증가시키고, 10% 이하일 경우에는 히스토그램의 분해수를 감소 시켜서 2번과정을 반복한다. 이 반복은 선택된 그룹들의 광류수가 10%에서 20% 사이에 오면 정지하게 된다. 이 알고리즘을 적용하는 과정에서 정합이 잘된 광류가 전체 광류에 대하여 40% 이상일 때는 알고리즘의 종료 기준을 10-20%보다 높은 값으로 설정해도 무관하나, 정합이 잘된 광류가 30% 이하일 경우나 영상회전과 줌이 있을 때는 정합이 잘된 광류만을 추출하기가 힘들다. 따라서 본 논문은 모든 영상에 대하여 정합이 잘된 광류를 추출하기 위하여 종료기준을 10-20%로 설정하였다. 또, 알고리즘을 2, 3번 반복 하기전에, 선택된 그룹의 전체 광류수가 20% 넘고, 광류수가 가장 많은 그룹이 전체 그룹들의 광류 수에 35% 이상일 때는 정합이 잘된 광류가 많다는 것을 의미하기 때문에 이 반복과정을 정지하고 광류수가 가장 많은 그룹만을 사용한다. 제한한 알고리즘에 의해 선택된 광류의 정확도를 평가 하기 위해 본 논문은 카메라와 화소간의



(a) 추출된 광류



(b) 방향과 길이로 구성된 광류



(c) 2차원 히스토그램

그림 3. 추출된 광류와 2차원 히스토그램

affine 식을 이용하였다. Affine을 계산하여 카메라 움직임에 의해 나타난 배경 광류를 구한후, Affine에 의해 추정된 위치와 다르게 움직이는 그룹들에서 광류가 3개 이상인 그룹을 선택하여 움직이는 물체로 인식하는 데 응용하였다.

### 3.1 2차원 히스토그램

본 알고리즘을 위한 2차원 히스토그램은 그룹화 확장 또는 축소영역과 고정영역으로 구성한다(그림 4 참조). 이렇게, 고정영역을 설정하는 이유는 길이가 작은 광류를 추출하기 위함인데, 길이가 작은 광류는 방향과 길이가 한정되어 있기 때문에, 그룹화 확장/축소를 진행하지 않는다. 이와는 달리 긴 광류 그룹은 방향과 길이가 다양하기 때문에 유사한 광류를 그룹화 하고, 그룹화된 광류수가 적으면 그룹화 확장, 그룹화된 광류수가 많으면 그룹화 축소한다.

이동량이 적은 카메라에의해 얻은 영상에서는 길이가 작은 광류들이 나타나는데, 이 광류들의 방향을 화소단위로 표현하는 경우 양자화되어 그 수가 한정되게 된다. 예를 들면, 길이가 1일 경우에는 45도, 2일 경우에는 22.5도, 3일 경우엔 15도, 4일 경우에는 11.25도 단위로 각각 8, 16, 32, 64개의 방향으로 표현되며, 길이에 따라 광류의 방향 수는 다음과 같다.

$$N_{\theta} = 2^{i+2}, \quad i=1,2,\dots, L_{max} \quad (7)$$

여기서,  $i$ 는 광류의 길이를 뜻한다. 그룹화 확장 또는 축소되는 영역은 고정영역과 다르게 광류의 길이와 방향이 많이 존재하기때문에 2차원 히스토그램의 분해수를 감소시키거나 증가시킬 필요가 있다. 이 분해수를 증감시키는 기준은 선택된 그룹의 광류수가 전체 정합된 수의 10%에서 20% 밖에 위치할 때 이다. 방향의 분해수는 최소 분해각  $\Delta\theta_{min}$ , 최대 분해각  $\Delta\theta_{max}$ 의 입력에 의해 결정된다.

$$N_{\theta}(L) = \frac{f(\Delta\theta_{max}) - f(\Delta\theta_{min})}{L_{min} - L_{max}} (L - L_{min}) + f(\Delta\theta_{max}) \quad (8)$$

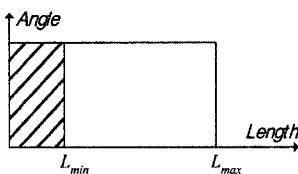


그림 4. 2차원 히스토그램에서 고정(빗금영역)과 변형영역.  $L_{min}$ 과  $L_{max}$ 은 각각 고정영역과 변형 영역의 최대 값

여기서, 2차원 히스토그램에서 고정영역과 그룹화 확장/축소영역 경계  $L_{min} = \frac{\ln(360/\Delta\theta_{max})}{\ln(2)} - 2$ 이며,  $f(\theta) = 360/\theta$ 이다. 식 (13)은 2차원 히스토그램의 가로축에서 길이가 증가할수록 방향의 분해수를 증가시키고 길이가 작아질수록 방향의 분해수를 감소 시키는 것을 의미한다. 계산된 결과는 소수점 이하를 제거한 정수로 양자화 되며, 그림 5와 같이 광류의 길이축 분해수는 최대 분해길이  $\Delta L_{max}$ 의 입력에 의해 식 (9)와 같이 결정된다.

$$N_L(\Delta L) = a_L(\Delta L - L_{min}) + 1 \quad (9)$$

여기서,  $a_L = \frac{\Delta L_{max} - 1}{L_{max} - L_{min}}$ 이다. 계산결과는 소수점 이하를 제거한 정수이다. 여기서, 분해를 위한 최소길이는 1로 한다.

$$N_{L,total} = \frac{\ln(\Delta L_{max})}{a_L} \quad (10)$$

식 (15)는 2차원 히스토그램의 가로축에서 길이가 증가할수록 길이 분해수를 감소시키고 길이가 작아질수록 길이 분해수를 증가 시키는 것을 의미한다. 물체 간의 거리가 가까운 물체에 대해서는 길이가 긴 광류가 추출되며 방향이 비슷하지만 길이변화가 많다. 반면에 투영된 물체가 카메라로부터 먼 거리에 있을 경우에는 길이 변화가 적다.

### 3.2 분해수 조정에 의한 그룹 수 증감

2차원 히스토그램의 분해수는 초기 설정값  $\Delta\theta_{min}$ ,  $\Delta\theta_{max}$ ,  $\Delta L_{max}$ 로 결정된다. 2차원 히스토그램에 정합된 광류를 인가하면 유사한 광류들로 그룹화가 되고, 광류수가 적은 그룹은 제거된다. 선택된 그룹들의 총 광류수가 정합된 광류수에 대하여 10%에서 20%에 존재하면 이 알고리즘을 종료한다. 또, 그룹화된 총 광류수가 20% 이상일 경우에는  $\Delta\theta_{min}$ ,  $\Delta\theta_{max}$ ,  $\Delta L_{max}$ 의 값을 감소시킬 필요

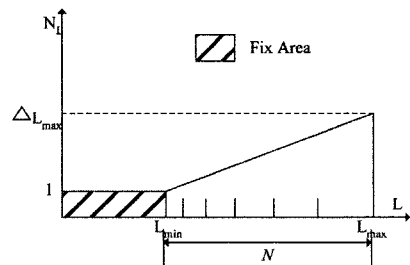


그림 5.  $\Delta L_{max}$ 에 대한 2차원 히스토그램의 가로축 분해수

가 있다. 이것은 2차원 히스토그램의 분해수를 증가시킴으로써 추출되는 광류수를 20% 미만으로 축소하는 효과를 얻을 수 있기 때문이다. 그룹화된 광류수가 전체 정합수의 20%를 초과하고, 동시에 선택된 그룹 가운데 가장 많은 비중을 차지하는 그룹이 나머지 그룹의 광류수의 35%를 초과하는 경우, 알고리즘 계산을 종료한다. 이때 가장 많은 비중을 차지하는 그룹만을 카메라 모델에 사용할 수 있다. 그룹화된 총 광류수가 10% 이하일 경우에는  $\Delta\theta_{\min}$ ,  $\Delta\theta_{\max}$ ,  $\Delta L_{\max}$ 의 값을 증가시킬 필요가 있다. 이것은 2차원 히스토그램의 분해수를 감소시킴으로써 추출되는 광류수를 10% 이상으로 확대하는 효과를 얻을 수 있기 때문이다. 이 과정에서 초기에 설정된 그룹들은 단일 그룹

화 되거나 여러 그룹으로 나누어지게 된다. 그림 6의 순서도는 이와 같은 제안한 알고리즘을 정리한 것이다. 그림 7은 2차원 히스토그램의 분해수를 확대/축소 할때, 2차원 히스토그램을 보여주고 있다.

#### 4. 실험결과 및 고찰

기존의 M-estimators와 LMedS 알고리즘을 사용하는 경우, 추출된 광류 오류가 각각 30% 혹은 35% 이상이 되면, 적용이 어렵다고 알려져있다[9,10]. 그런데, 회전과 줌이 클 경우나, 동일한 텍스처를 많이 포함한 영상경우 및 영상 번짐이 존재하는 경우에는 잘못 정합된 광류가 60% 이상 까지도 나타날 수 있기 때문에 본 알고리즘을 이와 같은 영상들에 적용실험하여 강건성을 확인하였다. 제안한 연구에서는 광류들을 카메라 모델에 의해 예측한 광류와의 차이 값에 따라 그룹화하여 오류 값이 가장 큰 광류 그룹을 제거하고, 광류 수가 일정한 수 이상인 그룹에 대해서는 상세 분할하는 과정을 반복하므로써 오류 광류의 비율이 매우 높은 경우에도 오류광류들을 제거할 수 있게하였다. 실험은 영상에서의 특징점 수를 1000개 이상 추출 하도록 SUSAN의 상수 G를 변화시켜 SUSAN 알고리즘을 반복하였다. 상관정합은 특징점을 중심으로 마스크 11을 사용하였다. 이 때, 정합 값이 0.8 이상인 특징점들만을 광류로 선정하였다. 또, 영상정보가 전혀 없는 것을 가정으로 하였기 때문에, 동일한 특징점을 찾기위한 영역은 영상전체로 하였다. 2차원 히스토그램에서 초기 상수인  $\Delta\theta_{\min}$ ,  $\Delta\theta_{\max}$ ,  $\Delta L_{\max}$ 을 각각 15도, 8도, 10픽셀로 설정 하였다. 그림 8,9,10은 영상회전과 줌, 카메라 이동, 유사한 텍스처가 많이 포함된 영상들에 대하여 각각 실험한 결과이다. 각 영상 위에 스크래치 모양의 선은 SUSAN operator를 이용하여 특징점 추출, 정합후 유사한 특징점들을 연결한 광류들을 뜻한다.

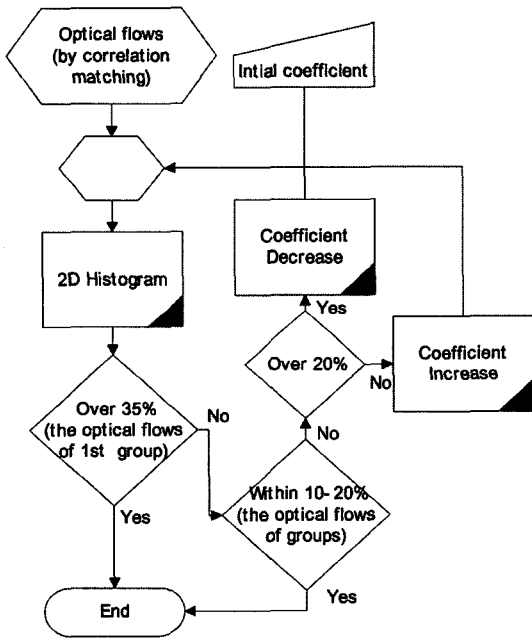
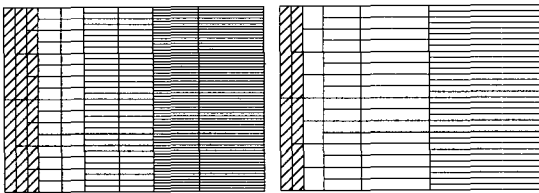


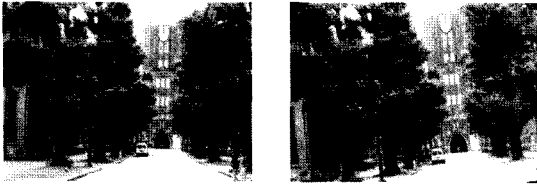
그림 6. 제안한 알고리즘 순서도



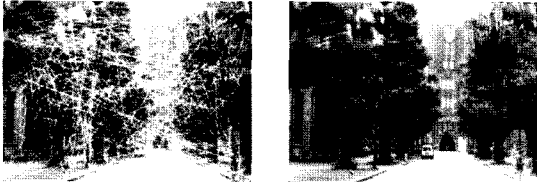
(a) 분해수가 많은 경우 (b) 분해수가 적은 경우

그림 7.  $\Delta\theta_{\min}$ ,  $\Delta\theta_{\max}$ ,  $\Delta L_{\max}$  값에 따른 2차원 히스토그램 분해수: 빗금친 영역은 고정영역을 뜻하며, 세로축은 방향, 가로축은 길이를 뜻한다.

그림 8(a)와 (b)는 카메라가 광축으로 회전과 줌 현상이 있는 영상이다. 그림 8(c)는 이 두 영상에서 특징점 추출후 상관정합 방법으로 계산하여 가장 유사성이 높은 특징점을 연결한 결과이다. 이 그림에서 보듯이 정합된 광류는 매우 복잡하고 불규칙한 결과를 보여주고 있다. 그림 8(d)는 제안한 알고리즘을 적용하여 그림 8(c)의 광류들로부터 유사성이 있는 광류를 추출한 결과인데, 지역적으로 유사한 광류가 그룹화되어 추출된 것을 알 수 있다. 그림 8(d)는 제안한 알고리즘으로 추출된 광류만을 이용하여 카메라와



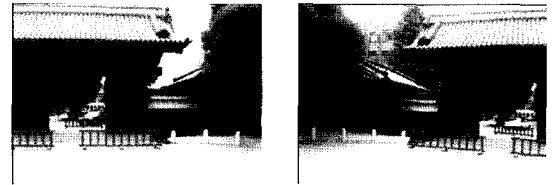
(a) 기준영상 (b) 다음 프레임 영상



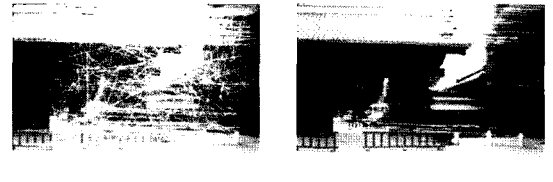
(c) 상관정합에 의한 광류 (d) 그룹화 후의 선택된 광류



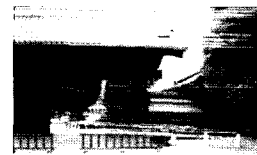
(e) 선택된 광류를 기반으로 affine 식에 의한 광류 재추출  
그림 8. 회전과 줌이 동시에 일어난 영상에서의 광류 추출



(a) 기준영상 (b) 다음 프레임 영상



(c) 상관정합에 의한 광류 (d) 그룹화 후의 선택된 광류

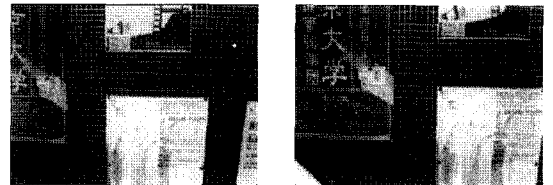


(e) 선택된 광류를 대상으로 affine 식에 의해 광류 재추출 결과  
그림 9. 영상 중첩이 40%인 영상에서의 광류

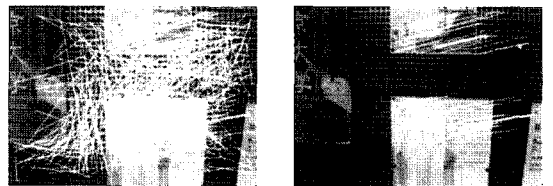
화소 간의 관계 식인 affine 식을 구한후, 다시 광류를 구했을 때의 결과이다.

그림 9(a)와 (b)는 중첩 영역이 40%인 영상이다. 그림 9(c)는 이 두 영상에서 특징점 추출후 상관정합 방법으로 계산하여 유사성이 가장 높은 특징점을 연결한 결과이다. 그림 9(c)에서 보는 바와같이 정합된 광류는 50%의 잘못 정합된 광류가 포함된 것을 보여주고 있다. 그림 9(d)는 제안한 알고리즘을 그림 9(c)의 광류들에 적용하여 유사성이 있는 광류들을 추출한 결과이다. 그림 9(e)는 제안한 알고리즘으로 추출된 광류들을 대상으로 affine을 구한후, 다시 광류를 구했을 때의 결과이다.

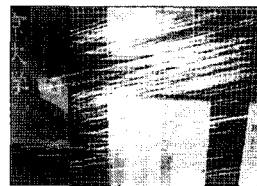
그림 10(a)와 (b)는 유사한 텍스처가 많이 포함된 영상이다. 그림 10(c)는 이 두영상에서 특징점 추출후 상관정합 방법으로 계산하여 가장 유사성이 높은 특징점을 연결한 결과이다. 그림 10(c)에서 보는 바와 같이, 정합된 광류는 매우 복잡한 결과를 보여주고 있다. 그림 10(d)는 제안한 알고리즘을 그림 10(c)에 적용 했을때, 유사성이 있는 광류를 추출한 결과이다. 그림 10(d)는 제안한 알고리즘으로 추출된 광류만을 이용하여 affine을 구한후, 다시 광류를 구했을때 결과이다.



(a) 기준영상 (b) 다음 프레임 영상



(c) 상관정합에 의한 광류 (d) 그룹화 후의 선택된 광류



(e) 선택된 광류를 대상으로 affine 식에 의해 광류 재추출 결과  
그림 10. 유사한 텍스처가 많이 포함된 영상에 대한 광류 추출 결과

표 1은 실험 결과의 요약으로써 모든 영상에 대하여 제안한 알고리즘은 강건하게 좋은 결과를 얻었음을 보여준다. 여기서, affine을 반복하여 광류수가 더 이상 증가하지 않을때, 이 광류들은 잘 정합된 광류라고 말할 수 있다. 초기 정합된 광류를 정합이 잘된 것과 정합이 잘못된 것을 평가하는 조건은 affine을 여러 번 반복하여 얻은 광류와 비교하여 일치하면 정합이 잘된 광류로 설정하고, 일치하지 않으면 정합이 잘못된 광류로 설정 하였다. 본 논문은 제안한 알고리즘으로 추출된 광류를 이용하여 affine을 1번 반복하여 얻은 광류들과 여러 번 반복하여 얻은 광류들과 비교 했을때, 동일한 광류를 비슷한 수로 얻을 수 있었다. 그림 8에서 초기 정합이 잘된 광류 수가 많았지만, 제안한 알고리즘에서 추출된 수가 적은 이유는 회전과 줌이

크게 발생하였기 때문이다. 이와 반대로 그림 9에서 많은 정합 광류들을 추출할 수 있었던 것은 영상의 이동 성분만을 포함하고 있기 때문이다. 표 2는 그림 8에서의 잘 정합된 광류를 기반으로 랜덤하게 발생시킨 광류를 추가한 후, LMedS와 제안한 알고리즘을 대상으로 정확한 광류를 추출하는 능력을 비교하였다. 실험에서 랜덤한 광류를 생성할때, 이 생성된 광류의 위치에 따라 결과가 조금씩 다르게 나오기 때문에 100번 실험한 결과를 평균하여 나타내었다. 랜덤 광류 발생 시 광류의 최대 길이는 입력영상의 가로 크기의 반으로 설정 하였다. 여기서의 LMedS알고리즘 실험을 위해서 affine식에 광류를 대입하여 모델을 구한후, 이 모델에 광류의 시작점을 입력하여 추정된 위치와 실제 광류의 끝점과의 거리가 어느 이상이 되면 광류오류로 간

표 1. 제안한 알고리즘을 그림 8, 9, 10에 적용한 결과 요약

그림 번호	전체수(inlier, %)	2차원 히스토그램			Affine반복 Iter1(inlier, %)/Iter5
		그룹수(반복수)	추출수(inlier, %)	추출수/전체수 %	
8	465(128, 26.9%)	4(2)	47(39, 82.9%)	10.1%	감소 217(217,100%)/217
9	426(49, 11.5%)	4(3)	67(57, 85.0%)	15.7%	상승 86(81, 94.1%)/96
10	407(36, 8.8%)	3(2)	41(36, 87.8%)	10%	감소 230(213, 92.6%)/233

Inlier: 맞는광류, outlier: 오류광류, Iter1: affine 1번 반복, Iter5: affine 5번 반복

표 2. 제안한 알고리즘과 LMedS 비교

outlier/inlier		Inlier	outlier	inlier/(outlier+inlier)	사용여부
0.6	LMedS	31	2	0.92	○
	Proposed	52	2	0.96	○
0.8	LMedS	18	3	0.82	○
	Proposed	51	1	0.97	○
1.0	LMedS	14	8	0.64	△
	Proposed	55	2	0.96	○
1.2	LMedS	13	10	0.56	×
	Proposed	51	1	0.97	○
1.4	LMedS	×	×	×	×
	Proposed	53	3	0.93	○
2.0	LMedS	×	×	×	×
	Proposed	54	5	0.90	○
2.8	LMedS	×	×	×	×
	Proposed	50	20	0.714	○
3.6	LMedS	×	×	×	×
	Proposed	45	29	0.606	△
4.0	LMedS	×	×	×	×
	Proposed	38	41	0.48	×

Inlier: 맞는광류, outlier: 오류광류 ○: 맞는광류 추출 가능성 높음, △: 맞는광류 추출 가능성 있음, ×: 맞는 광류 추출 불가능



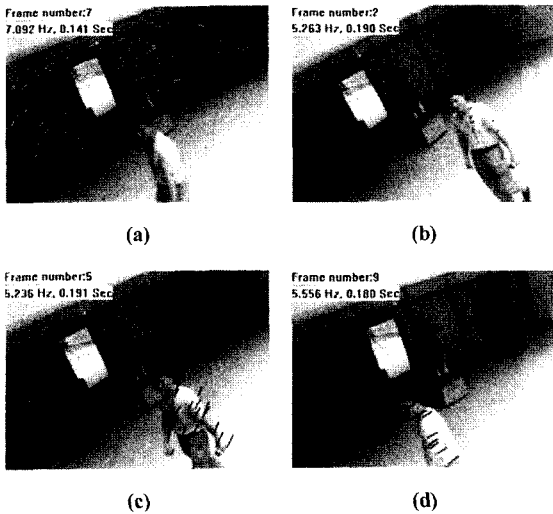


그림 11. 회전하는 카메라로 촬영한 영상에서 움직이는 물체 추적 실험 결과: (a) 배경 광류, (b), (c), 및 (d) 움직이는 물체들에 대한 광류

주하여 그 광류를 제거하는 방식을 사용하였다. 이 실험에서 광류오류는 20픽셀로 설정 하였다. 표 2에서 보듯이 제안한 알고리즘은 랜덤하게 생성된 광류수가 잘 정합된 광류의 수의 4.0배 가까이 되어도 강인하게 잘 정합된 광류를 추출해냄을 보이고 있다. 이와 반면에 LMedS은 랜덤하게 생성된 광류수가 잘 정합된 광류의 수의 1.0배 이상일 때는 계산이 불가능 하였다. 그림 11은 움직이는 카메라에서 움직이는 물체의 추적 결과로서 320\*240인 영상에 대해 제안한 알고리즘을 이용한 실험한 결과이다. 이 실험에서 특징점 추출/매칭과 제안한 알고리즘 적용을 한 사이클로 했을때, 팬티엄 II(220MHz)를 사용하여 초당 5프레임을 계산할 수 있었다. 그림 11(a)는 제안한 알고리즘을 적용 후, affine을 계산하여 카메라 움직임에 의해 나타난 배경 광류를 구한 결과이고, 그림 11(b), (c) 및 (d)는 제거된 그룹들에서 광류가 3개 이상인 그룹을 선택한 결과이다.

그림 11에서 보는 바와 같이 제안한 알고리즘은 카메라가 회전/이동하면서 영상 내에서 움직이는 물체를 빠르고 정확하게 추적하고 있음을 보여준다. 움직이는 물체의 광류를 사용하여 지역 영상분류를 한다면, 효율적으로 빠르게 움직이는 물체의 경계를 추출 할 수도 있게된다.

## 5. 결 론

본 논문은 광류의 오류율이 85% 이상인 경우에도 강건

하게 정합이 잘된 광류를 찾는 가변 분해능을 가지는 2차원 히스토그램 이용방법을 제안하였다. 2차원 히스토그램은 유사한 광류를 그룹화하는 것으로서, 2차원 히스토그램을 분해수를 변화시켜 정합이 잘된 광류수가 전체수에 대하여 10%에서 20%사이에 오도록 선택하는 알고리즘이다. 제안한 알고리즘을 이용한 실험결과 광류 오류율이 높은 다양한 종류의 영상들에 대해서 잘된 광류들만을 추출할 수 있는 좋은 성능을 확인할 수 있었다.

outlier제거 후, RANSAC(Random Sample Consensus) 방법으로 block adjustment을 이용하여 카메라 오리엔테이션 계산할 때의 계산소요가 비약적으로 감소 될 것으로 기대된다. 카메라가 움직이면서 물체가 움직일 때, 제거된 outlier중에서 움직이는 물체를 인식해 낼 수 있으므로 자율주행 로봇등에 응용될 수 있을 것으로 기대된다.

## 참고문헌

- 이진덕, 최용진 (2001), CCTV유형 CCD 카메라를 이용한 근거리 산업사진측량의 정확도, 한국측량학회지, 제 19권, 제 3호, pp. 283-290.
- 장호식, 서동주, 이종출 (2003), 비디오 영상을 이용한 석조 문화재 위치해석, 한국측량학회지, 제 21권, 제 4호, pp. 355-363.
- Adelson, E. H. and J. R. Bergen, 1985, Spatiotemporal energy models for the perception of motion, *Journal of Optical Society of America, A*, Vol. 2, No. 2, pp. 284-299.
- Haralick, R.M. et al., 1988, Pose estimation from corresponding point data. *IEEE Tr. SMC*, 19(6):1426-1446, Nov.
- Harris, G., 1987, Determination of Ego-Motion From Matched Points, *Proc. Alvey Vision Conf.*, Cambridge UK.
- Horn BKP, Schunck BG, 1981, Determining Optical Flow, *Artificial Intelligence* 1981, pp. 185-203.
- Kearney, J. K., W. B. Thompson, and D. L. Boley, 1987, Optical flow estimation: An error analysis of Gradient-based methods with local optimization, *IEEE Tr. on PAMI*, Vol. 9, No. 2, pp. 229-244.
- Kitchen, L., and A. Rosenfeld, 1982, Gray Level Corner Detection, *Pattern Recognition Letters*, pp. 95-102.
- Mokhtarianm, Farzin, and Riku Suomela, 1998, Robust Image Corner Detection Through Curvature Scale Space, *IEEE Tr. on PAMI*, Vol. 12.
- Rousseeuw, P. J., 1984, Least median of squares regression. *Journal of American Statistics Association*, 79:871-880.
- Smith, S.M. and J. M. Brady, 1995, Real-Time Motion Segmentation and Shape Tracking, *IEEE Tr. on PAMI*, Vol. 17, No. 8.
- Smith, S.M., and J.M. Brady, 1997. SUSAN - a new approach to low level image processing. In *IJCV*, 23(1), pp. 45-78.

(접수일 2005. 2. 20, 심사일 2005. 2. 28, 심사완료일 2005. 3. 26)