공학 연구논문집 제26권 제1호 pp. 275~289, 울산대학교
Journal of Engineering Research Vol. 26, No. 1, pp. 275~289 1995.
University of Ulsan.

영역의 계층 구조화에 의한 이동 변위 추정

김만우 · 최원호
전자공학과

<요 약>

본 논문에서는 물체의 움직임이 있는 영상에 대하여, 밝기 함수의 공간 및 시간적 기울기를 토대로한 이동 변위 추정 방법에서 기존 알고리즘의 문제점을 해석하고, 이를 개선시키는 알고리즘에 대해서 논하고자 한다.

기울기를 토대로 하는 기존의 방법은 국부적인 변위를 추정하므로 전체적인 광역정보를 상실하여 오차를 발생한다. 그리고 이동 물체의 밝기 변화가 균일한 영역에서는 불완전한 조건(ill-condition)과 부적절한 배치(ill-posed)로 인한 구경(aperture) 문제로 인하여 이동 속도 추정이 어렵다. 이와 같은 문제를 해결하기 위해 운동 제한선을 이용하여 선형적으로 이동 속도를 구하고, 이를 근거로 계층구조를 구축한다. 선형적으로 구한 속도들을 4×4 윈도우로 표본화하여 확률 밀도 함수를 구하고, 확률 밀도 함수가 가장 큰 속도를 계층 구조의 초기 값으로 설정하여 광역으로 전파시킨다. 계층 구조로 형성한 결과 기존의 구경(aperture) 문제를 개선시키고, 정상 상태 오차를 줄이고, 수렴 속도를 개선시켰다.

An Optical Flow Estimation Method By The Hierarchical Structuring of Image Segments

Man Woo Kim · Won Ho Choi
Dept. of Electronic Engineering

<Abstract>

The relationship between the gradients of brightness of image is an important factor for optical flow estimation techniques. Gradient-based methods are sensitive to conditions commonly encountered in real images. Highly textured surface, large constant brightness area, motion boundary and depth discontinuity can all be troublesome. These factors can be a cause of aperture problems, flow discontinuities, and so on.

In this paper, the sources of errors of local gradient-based techniques are analyzed, and an optical flow estimation algorithm using hierarchical structuring of image segments is proposed.

With the hierarchical structure the effective combination of local and global information can be achieved. The algorithm uses the image pyramids those are from the repetitive application of Gaussian filtering and decimation of the 2^n:1. For faster convergence and lower steady state errors, we used intersections of constraint lines as initial values of the first pyramid level. The experimental results on various data reveal that the proposed algorithms show faster convergence and more accurate estimation results than conventional methods.

Ⅰ. 서 론

인간의 시각인식은 인간의 사고와 인식에 있어 커다란 비중을 차지한다. 그리고 배경에 대한 물체의 연속적인 움직임은 인간의 인식과 판단에 지대한 영향을 미친다고 알려져 있다[1]. 물체의 움직임이 있는 연속 영상으로부터 물체 운동의 해석은 물체의 구조나 3차원 운동을 이해하는데 도움을 준다. 그 응용으로써 위성추적, 미사일의 목표물 탐사, 움직이는 물체의 자동 추적, 비행기의 자동 이착륙, 일반적인 영상과는 다른 레이더, 수중 음파 탐지, 적외선 영상으로부터 움직이는 물체의 추출, 무인 자동차의 움직이는 방향과 속도에 대한 정보 수집 등이 있다. 또한, 의학, 기상학, 미생물 활동 관찰, 동화상 압축 등에도 응용되어지고 있다. 이와 같은 다양한 응용으로 인하여 움직이는 영상에 대한 연구가 활발하게 이루어지고 있다.

물체의 움직임을 추출하기 위한 시변 영상(time varying image)의 분석은 크게 영상 데이타로부터 물체의 움직임에 대한 정보를 추출하는 단계로부터 시작된다. 우선 운동 영상에 대한 속도 변위를 추출해야 한다. 이동 변위 추출은 움직이는 물체의 특징을 근거로 하는 방법과 밝기를 근거로 하는 방법 등이 있다.

특징을 근거로 하는 방법은 연속적인 두 영상의 특징을 추출하여 정합시킴으로써 움직이는 물체의 변위를 얻는다. 물체의 밝기를 근거로 하는 방법은 시간에 따른 밝기 함수의 움직임과 밝기의 지역적 구조로부터 이동 변위(optical flow)를 얻을 수 있다. 각 화소의 이동 변위가 구해지면 이를 근거로 하여 움직이는 물체를 분할하거나 3차원 운동[4,5]을 해석할 수 있다. 그런데, 이 두 과정은 수학적인 관점에서 문제점이 존재한다. 즉, 카메라의 이동이나 물체의 이동은 실제에서는 3차원적 움직임이지만 영상데이타에서는 투영된 2차원 데이타이므로 깊이 정보를 손실한다. 이로 인해 먼저 구경(aperture) 문제가 발생한다. 한 화소와 이웃 화소간의 밝기의 기울기와 연속영상 간의 밝기 차로 이동 변위를 추정하므로 전체적인 움직임과는 별도로 2차원 변위 대신 하나의 성분만 추정하게 된다[6]. 또 하나는 영상내의 물체 움직임이 2차원인 영상데이타의 밝기 변화에 전적으로 의존하지 않게 되어, 영상내의 물체가 움직였음에도 불구하고 물체의 밝기가 균일한 부분은 변위를 구할 수 없게 된다[6].

본 논문에서는 전술한 물체의 밝기를 기초로 한 이동 변위 추정에 관하여 논하고자 한다. 이 방법으로 물체의 이동 변위를 구하기 위해서는 공간 경사와 시간 경사를

표현하는 제한식 외에 영상의 밝기 변화가 공간상에서 크게 변하지 않는다는 추가적인 제한이 필요하다[6]. 이 제한으로 인하여 이동 변위가 큰 부분에서는 이동 벡터가 잘못 구해지기도 하고, 움직이는 물체의 균일한 밝기 부분에서는 기울기가 일정하므로 변위 구하기 매우 힘들다.

이 문제를 해결하기 위해 많은 연구가 이루어져 왔다. Yachida[8]는 모서리와 같은 특징적인 부분에서 변위를 구한 다음 이를 Horn과 Shunk[6]가 제안한 반복적 접근방식을 이용하여 속도를 전체로 전파시켰다. Wu[9]는 모서리에서의 속도는 추가적인 제한없이 속도를 구할 수 있다는 점을 이용하여 모서리에서 속도를 구하여 경계선을 따라 추정된 속도를 전파시켰다. Cornelius와 Kanade[10]는 움직이는 경계 부근에서 완만성의 정도를 약하게 하여 이동 변위가 흐릿해지는 것을 막았다. Shunk는 속도의 크기와 경사의 방향으로 구성된 운동제한식을 제안하고, 제한선을 토대로 무리화하는, 제한선 무리화(Constraint Line Clustering, CLC)를 이용하여 운동경계와 이동 변위를 동시에 추정하는 방법을 제안하였다[11]. 또한, 공간 경사나 시간 경사는 미분 개념이므로 실제 물체의 이동과는 오차가 발생하는데, Kearney 등은 국부적 최적화(local optimization)를 이용하여 이동 변위를 구할 때 경사들의 오차를 분석하였다[13]. 그리고, Nagel 등은 측정된 오차 중에서 움직이는 물체의 밝기 변화의 비선형성 때문에 발생하는 오차를 줄이기 위해서 이동된 프레임의 차(Displaced Frame Difference, DFD)를 구하여 광역적 최적화(global optimization)하는 반복 접근식을 사용하였다.

본 논문에서는 국부 정보의 체계적 통합과 국부 정보와 광역 정보를 효율적으로 결합하는 계층적 움직임 추정 방법(hierarchical motion estimation method)을 사용하여 구경 문제를 해결하였다. 그리고 경계부분의 흐릿해짐을 최소화하고, 반복 접근식의 수렴 속도를 높이기 위해 4×4 윈도우 내의 제한선의 교점들 중 확률 분포 함수가 가장 높은 점을 최초의 계층구조의 초기값으로 사용하여 계층구조를 통해 전체적으로 전파시켰다.

Ⅱ. 이동 변위 추정

1. 기울기를 기초로 한 이동 변위 추정

연속적인 영상간의 밝기 차이, 밝기의 변화를 이용하여 움직임 정보를 비교적 간단한 계산으로 구할 수 있는 밝기의 시간 경사(temporal gradient) 및 공간 경사(spacial gradient)를 이용한 변위 추정 방법이 이동 변위 계산에 광범위하게 사용되고 있다. 그리고 이동 변위 추정시 다음 사항들이 가정되어야한다. 첫째 영상의 조명 상태는 항상 일정하다. 둘째 한 점의 밝기 변화는 오로지 물체의 움직임으로부터 비롯된다. 셋째 물체의 움직임에 따른 밝기 변화는 완만하다. 넷째 영상 데이타는 수학적으로 불연속성을 갖지 않는다.

임의 시간 t에서 (x,y)점의 밝기 함수를 I(x,y,t)라고 하자. 만약, Δt시간 후에 (x, y)점이 $\Delta x, \Delta y$만큼 이동했다고 하면,

$$I(x, y, t) = I(x+\Delta x, y+\Delta y, t+\Delta t) \tag{2-1}$$

이 된다. 우변을 Taylor급수로 전개하여 고차항을 무시하면 운동 제한식을 얻을 수 있다.

$$I_x u + I_y v + I_t \cong 0 \tag{2-2}$$

$$\left(u = \frac{\Delta x}{\Delta t},\ v = \frac{\Delta y}{\Delta t}\right)$$

물체가 움직인 국부 속도는 제한선 위에 존재한다. 움직인 속도와 방향은 제한

선으로부터 구할 수 있다(그림 2-1).

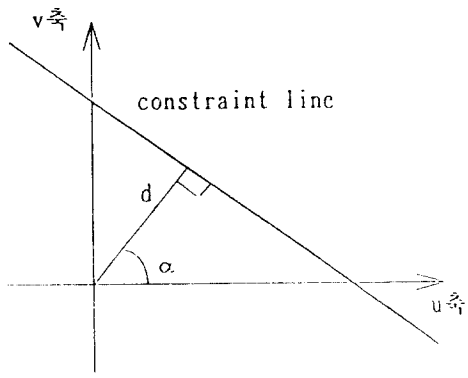


그림 2-1 제한선
Fig. 2-1 constraint line

그림 2-1에서 d는 물체가 움직인 거리이고, α는 움직인 방향이다.

$$d = -\frac{I_t}{\sqrt{I_x^2 + I_y^2}} \tag{2-3}$$

$$\alpha = \tan^{-1}\left(\frac{I_y}{I_x}\right) \tag{2-4}$$

그런데 독립된 한 점의 밝기만으로는 물체의 속도가 유일하게 결정되지 않는다. 즉, 한 물체의 이동은 이동 물체의 모든 점에서 같은 속도를 가지므로 물체의 한 점에서 속도가 독립적으로 결정되지 않는다[6,13]. Horn Shunk는 속도를 전체적으로 구하기 위해, 움직이는 물체의 이웃 점들의 밝기 함수의 변위장(flow field)에서는 속도가 완만하게 변한다는 완만성에 대한 제한(smoothness constraint)을 제안하였다[6]. 이동 변위 속도 기울기의 제곱크기(식2-5)를 최소화시키는 것이 또 다른 방법이다.

$$\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 \text{ 그리고 } \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 \tag{2-5}$$

u,v를 라플라시안으로 표현하면 다음과 같다.

$$\Delta^2 u = \frac{\partial^2 u}{\partial x^2} \text{ 그리고 } \Delta^2 v = \frac{\partial^2 v}{\partial y^2} \tag{2-6}$$

위 식을 디지탈화하기 위해 그림 2-2의 템플릿을 사용하여 근사화시킬 수 있다.

$$\Delta^2 u \approx k(\bar{u}_{ijk} - u_{ijk}) \text{와 } \Delta^2 v \approx k(\bar{v}_{ijk} - v_{ijk}) \tag{2-7}$$

(k = 비례 상수)

$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$
$\frac{1}{6}$	-1	$\frac{1}{6}$
$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$

그림 2-2 라플라시안 템플릿
Fig. 2-2 Laplacian Template

전술한 두 식을 가중치로 합하여 적분값을 최소화하는 전체적인 최적화 방법을 이용하면 이동 변위를 예측할 수 있다.

$$\int\int (I_x u + I_y v + I_t)^2 + \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}\right)^2 dxdy \tag{2-8}$$

여기서 가장 알맞은 이동 변위(u,v)를 구하면 최적화시킬 수 있다[6]. 전체적으로 최적화하기 위해서 Gauss-Seidal 방법으로 반복적 해답을 구하면 다음과 같이 된다.

$$u^{n+1} = \bar{u}^n - \frac{I_x[\ I_x\bar{u}+I_y\bar{v}+I_t]}{(\alpha^2+I_x^2+I_y^2)} \tag{2-9}$$

$$v^{n+1} = \bar{v}^n - \frac{I_y[\ I_x\bar{u}+I_y\bar{v}+I_t]}{(\alpha^2+I_x^2+I_y^2)} \tag{2-10}$$

2. 오차 해석

2.1 기울기 측정 오차

밝기 함수의 기울기를 토대로 한 방법은, 실제 영상에서 나타나는 밝기가 일정한 영역, 복잡한 무늬영상, 운동경계와 깊이정보 등에 매우 민감하며, 공간경사 I_x, I_y와 시간경사 I_t는 공간적, 시간적 표본화 간격에 따라 이들의 영향 정도가 결정된다. 밝기 함수의 랜덤한 오차는 채널 잡음, 양자화 정도와 프레임 간격과 같은 시스템 성능에 관련된 오차이다. 비록 영상에 잡음이 없다하더라도 공간적, 시간적으로 표본화되어 있으므로 실제 영상의 한 화소가 표본 간격의 중간 위치에 존재하면 가까운 화소로 근사화될 수 밖에 없다. x방향의 표본화 간격이 Δx라하고, x방향의 공간 경사를 구하기 위해 임의의 (x, y,t)점이 Δx만큼 이동되었다고 가정하면 다음과 같이된다.

$$I(x+\Delta x, y, t) = I(x, y, t) + I_x\Delta x - \frac{I_{xx}}{2}\Delta x^2 + h.o.t \tag{2-11}$$

(I_x, I_{xx}: (x, y, t)에서 편미분)

x 방향의 공간 경사로 다음과 같은 근사값을 사용할 수 있다.

$$\hat{I_x} = \frac{I(x+\Delta x, y, t)-I(x, y, t)}{\Delta x} = I_x + \frac{I_{xx}}{2}\Delta x + h.o.t \tag{2-12}$$

h.o.t는 근사적으로 무시하면 x방향의 표본화로 인한 기울기 측정 오차는 다음과 같다.

$$\varepsilon_{x(\text{sampling})} = \hat{I_x} - I_x \approx \frac{I_{xx}}{2}\Delta x \tag{2-13}$$

마찬가지로 y,t 방향의 표본화 오차를 구할 수 있다.

$$\varepsilon_{y(\text{sampling})} \approx \frac{I_{yy}}{2}\Delta y \tag{2-14}$$

$$\varepsilon_{t(\text{sampling})} \approx \frac{I_{tt}}{2}\Delta t \tag{2-15}$$

2.2 변위장의 불일치 오차

변위장에서 실제의 물체의 이동 방향은 거의 균일하게 나타난다. 제한 조건 등에 의해서 왜곡될 경우, 실제 같은 방향임에도 불구하고 다르게 나타난다. 점(x_i,y_i,t_i)에서 변위가 가까운 이웃 점(x_j,y_j,t_j)을 고려해보자.

$$I_x^{(i)}u + I_y^{(i)}v + I_t^{(i)} = 0 \tag{2-16}$$

$$I_x^{(j)}u+I_y^{(j)}v+I_t^{(j)}=0 \tag{2-17}$$

불일치를 고려하면 오차를 구할 수 있다.

$$I_x^{(i)}u+I_y^{(i)}v+I_t^{(i)}=0 \tag{2-18}$$

$$I_x^{(j)}(u+\Delta u)+I_y^{(j)}(v+\Delta v)+I_t^{(j)}=0$$

$$I_x^{(j)}u+I_y^{(j)}=-I_t^{(j)}-\varepsilon_{\Delta\text{flow}} \tag{2-19}$$

여기서 $\varepsilon_{\Delta\text{flow}}$는 회전과 같은 불균등한 변위 발생 시 오차가 크게 나타난다.

2.3 조건[15]

이 문제는 영상 밝기 함수의 비선형성으로부터 오는 오차로서 자연 영상에서 많이 발생한다. 영상의 기울기가 선형이면 정확한 속도 (u,v)를 구할 수 있다. 그러나 기울기가 조금만 변하면 속도가 크게 달라지게 되는데 이를 불완전한 조건(ill-conditioned)이라고 한다.

만약 식(2-16), (2-17)을 벡터로 표현하면 다음과 같다.

$$\vec{G}\cdot\vec{w}=-\vec{b} \tag{2-20}$$

$$\vec{G}=\begin{bmatrix} I_x^{(i)} & I_y^{(i)} \\ I_x^{(j)} & I_y^{(j)} \end{bmatrix},$$

$$\vec{w}=\begin{bmatrix} u \\ v \end{bmatrix}, \quad \vec{b}=\begin{bmatrix} I_t^{(i)} \\ I_t^{(j)} \end{bmatrix}$$

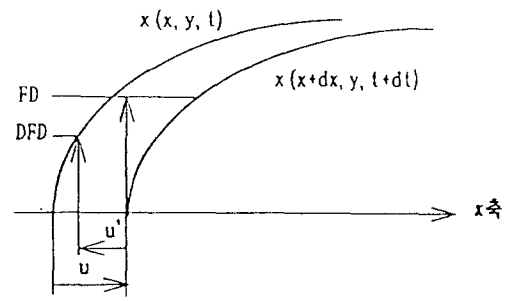
가장 널리 쓰는 조건식은 [23]에서 제안한 식이다.

$$\text{cond}(\vec{G})=\|\vec{G}\|\,\|\vec{G}^{-1}\|=\frac{\|\vec{g}^{(i)}\|^2\,\|\vec{g}^{(j)}\|^2}{\|\vec{g}^{(i)}\|\,\|\vec{g}^{(j)}\|\sin\Phi} \tag{2-21}$$

cond()에서 가장 중요한 것은 제한선들 사이의 각 Φ이다. 만약 Φ가 0이면 제한식은 서로 겹치므로 실제의 해를 구할 수 없다. 이로 인해 발생하는 문제가 구경 문제이다.

3. 전체의 최적화

시간 경사와 공간 경사로부터 속도 (u,v)를 구할 때 두 개의 식이 필요하다. 첫번째 식은 한 화소의 속도를 의미하고, 이를 이동 물체 전역으로 확장시키기 위해 추가적인 식(smoothness constraint)을 많이 사용한다[4,13]. 이 두 식을 최소화하는 속도 (u,v)를 구하면 된다. 이것은 식(2-8)을 사용하여 구한다. Horn과 Shunk는 식(2-9)과 같은 반복적 접근식을 사용하여 전체적으로 전파시켰다. Nagel과 Enkelman은 표본 오차를 고려하여 시간 경사 DFD(Displaced Frame Difference)[6]를 구하였다.



u : 실제 이동 속도
dx : x-축 방향의 이동거리
u' : 추측된 이동거리

그림 2-3 밝기 함수 이동에 대한 시간 경사
Fig. 2-3 The temporal gradient for the motion of the brightness function

FD와 DFD의 반복적 접근식을 비교하면 다음과 같다.

(1) Horn과 Shunk (FD)

$$u^{n+1} = \bar{u}^n - \frac{I_x(I_x \bar{u}^n + I_y \bar{v}^n + FD)}{\alpha^2 + I_x^2 + I_y^2} \quad (2\text{-}22)$$

$$v^{n+1} = \bar{v}^n - \frac{I_y(I_x \bar{u}^n + I_y \bar{v}^n + FD)}{\alpha^2 + I_x^2 + I_y^2} \quad (2\text{-}23)$$

(2) Nagel과 Enkelman (DFD)

$$u^{n+1} = \bar{u}^n - \frac{I_x(I_x \bar{u}^n + I_y \bar{v}^n + DFD)}{\alpha^2 + I_x^2 + I_y^2} \quad (2\text{-}24)$$

$$v^{n+1} = \bar{v}^n - \frac{I_y(I_x \bar{u}^n + I_y \bar{v}^n + DFD)}{\alpha^2 + I_x^2 + I_y^2} \quad (2\text{-}25)$$

원 이미지에는 고주파 성분이 포함되어 있기 때문에 전체적으로 불규칙적이고, 고주파 경계부분의 불완전한 조건때문에 잘못된 이동 속도가 구해진다. 이 영향을 줄이기 위해 가우시안 필터를 사용하여 고주파의 영향을 줄였다. 가우시안 필터의 식은 다음과 같다.

Gaussian Filter :

$$G = \frac{1}{64}\begin{bmatrix} 1 & 3 & 3 & 1 \\ 3 & 9 & 9 & 3 \\ 3 & 9 & 9 & 3 \\ 1 & 3 & 3 & 1 \end{bmatrix} \quad (2\text{-}26)$$

Ⅲ. 계층구조를 이용한 변위추정 알고리즘

Ⅱ.2절에서 논의한 오차문제 중에 불완전한 조건 문제로 인한 구경 문제를 해결하고 밝기가 일정한 물체 내의 변위를 찾기 위해 많은 연구가 이루어지고 있다. 이러한 문제점들을 해결하기 위해 국부 정보를 광역 정보로 체계적으로 결합하는 방법으로 계층적 움직임 추정에 대한 연구가 많이 이루어지고 있다. 정신 물리학 측면에서 현재까지 연구된 결과를 살펴보면 인간의 인식구조는 물리적 및 정신적 계층형태를 이루고 있음이 밝혀졌다[16]. 특히, 물체의 관측시 광역 정보로 전체적인 윤곽을 인식하고 점차 국부 정보인 물체의 부분적인 세부사항을 인식함으로써 구경 문제를 극복함을 알 수 있다. 이러한 관점에서 볼때 컴퓨터를 사용한 변위 추정에서는 각 계층으로 나누어 각 계층별로 특징있는 정보를 추출하여 변위를 구하고, 낮은 계층에서 높은 계층으로 전파시켜 효율적인 변위를 추정할 수 있다는 것이다. 영상 피라미드 구성 시에 각 단계의 표본 윈도우의 크기를 잘 고려해야 하는데 보통 그림 3-1과 같이 규칙적인 상태로 만들어 피라미드를 구성한다. 각 단계의 간축율(decimation rate)은 2^n:1을 보통 사용한다.

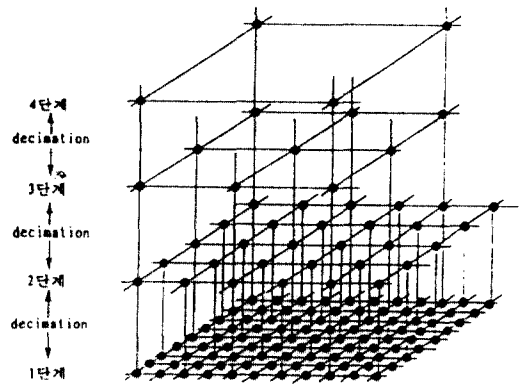


그림 3-1 피라미드 구조
Fig. 3-1 pyramid structure

영상 피라미드 구성 시 또한 중요한 문제 중 하나는 필터의 선택 문제를 들 수 있다. 필터는 각 피라미드 간의 대역제한을 위해 사용한다. 간축이 되면 영상의 불연속성때문에 고주파가 생성되어 대역제한 없이는 왜곡된 정보를 얻기 쉽다. 일반적으로 주파수 특성을 좋게 하기 위해서는 필터의 커넬(kernel)의 크기가 커야하는 반면 계산량과 하드웨어 구성 시 문제점을 적절히 고려하여 결정한다. 이러한 주파수 응답과 커넬 크기의 절충은 불확정성(un certainty relation)[17]에 의해 정량적으로 이루어진다. 이 관계에 따르면 가우시안 함수가 저역 통과 필터로서 절충효과를 보이며[18] 회전 대칭성을 보유하고 있다. 또한 각 단계별로 커넬을 따로 사용하면 비효율적이므로 Burt[19]는 저역 통과 필터링 및 간축을 반복적으로 적용하여 수행하였다. 각 층의 이동 변위를 구하기 위해 반복적 접근 방법으로 앞에서 설명한 FD, DFD 등의 알고리즘을 사용한다.

반복적 접근 방법을 사용하여 상위층의 결과를 추정한 다음 이 결과를 다음 하위층의 초기값으로 사용하여 계속해서 그 단계의 전체로 전파시킨다. 오차가 한계치 이하로 줄어들 때까지 반복한다. 이렇게 하면 광역의 정보가 전 영상에 전파되어 한 화소의 공간적 시간적 기울기를 이용한 FD, DFD 등의 추정보다는 전체적인 변위 추정이 유리하다는 것이다. 이 알고리즘을 N×N 영상 데이타에 대해 L계층의 계층적 변위 추정 알고리즘에 적용하면 다음과 같다.

알고리즘

단계 1. Let $u^{proj} = v = 0$; k=L

단계 2. Calculate

$$P = I_x^{proj} + I_y^{proj} + I_t$$

$$D = \alpha^2 + I_x + I_y$$

$$u^k = u^{proj} - I_x \frac{P}{D}$$

$$v^k = v^{proj} - I_y \frac{P}{D}$$

단계 3. $u^{proj} = P_k^{k-1}(u^k)$, where $P_n^m(\ )$ is a projection n operator
$v^{proj} = P_k^{k-1}(v^k)$ of level n into m
k = k - 1

단계 4. Calculate

$$P = I_x^{proj} + I_y^{proj} + I_t$$

$$D = \alpha^2 + I_x + I_y$$

$$u^k = u^{proj} - I_x \frac{P}{D}$$

$$v^k = v^{proj} - I_y \frac{P}{D}$$

단계 5. $u^{proj} = P_k^{k-1}(u^k)$
$v^{proj} = P_k^{k-1}(v^k)$
k = k - 1

단계 6. If (converge) k = k - 1
If (k = 1) stop
else goto step 2

Ⅳ. 제한선을 이용한 계층 구조의 변위 추정

FD, DFD와 같이 공간 경사와 시간 경사를 이용하여 한 화소의 국부적인 변위를 추정할 때 나타나는 불완전한 조건, 구경, 부적절한 배치 등의 문제를 해결하기 위해 인간의 시각 구조와 비슷한 계층 구조를 이용하여 이동 속도를 추정한다. 저해상도의 추정 속도를 고해상도의 반복적 접근식의 초기값으로 사용하여 저해상도의 이동속도가 전체적으로 전파되어 가도록 한다.

고해상도에서 저해상도로 표본화할 때 물체의 이동 속도를 유지하고 표본화로 인한 영상의 불연속성과 잡음을 막기 위해 적절한 저역 통과 필터가 필요하다. 피라미드 계층이 많을수록 큰 변위 추정에 이점이 있지만 한 계층이 추가될 때마다 간축과 필터링을 거치게 되므로 그림 4-1과 같이 이동 물체의 윤곽선과 같은 고주파 성분에 대한 손실이 발생하게 된다. 저해상도 피라미드 계층으로 갈수록 피라미드 표본화(그림 3-1)에 따른 랜덤 잡음이 발생하여 심하면 물체의 움직임과는 전혀 다른 변위를 추정할 수 있다. 계층적 표본화 오차는 계층에 따라 누적되어 증가한다.

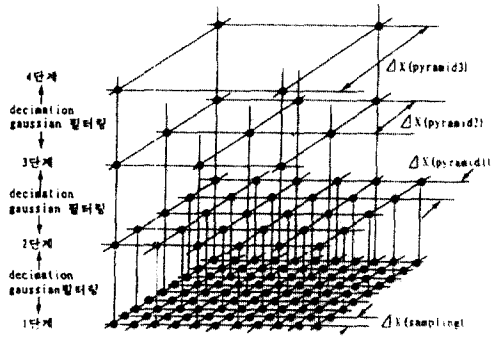


그림 4-1 피라미드 계층의 표본화 및 필터링

Fig. 4-1 Sampling and filtering of pyramid structure

본 논문에서는 앞서 해석한 윈도우 표본화로부터 비롯된 오차를 줄이고 계층구조 변위 추정의 특성을 유지하면서 이동 속도를 추정하기 위해 선형적인 운동 제한식의 해를 구하였다. 그리고 실제 영상의 비선형성을 고려하기 위해 Horn과 Shunk가 제안한 반복적 접근 방법으로 2단계 피라미드 계층 구조를 통하여 전체적으로 전파시키므로써 비선형으로 인한 오차를 줄였다.

1. 제한선

운동 제한선 모델, 식(2-2)을 극좌표 형태로 나타내면 다음과 같다.

$$\rho\,|\sqrt{I_x^2+I_y^2}|\cos(\alpha-\beta)$$

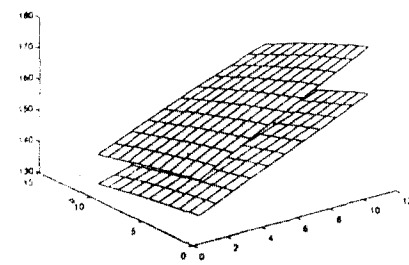
$$\left(\begin{array}{c}\beta:\text{물체의 움직이는 방향}\\ \rho:\text{움직이는 속도}\end{array}\right) \qquad (4\text{-}1)$$

$\sqrt{I_x^2+I_y^2}$로 나누면,

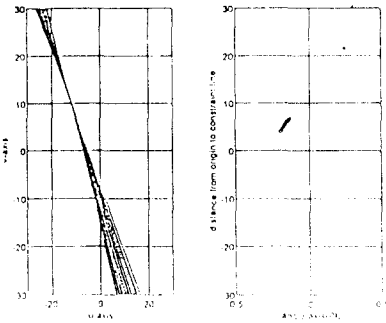
$$d=\rho\cos(\alpha-\beta) \qquad (4\text{-}2)$$

가 된다. d는 항상 절대값이고, 속도 (u,v)는 각 α의 기울기를 갖는 d를 u축과 v축에 각각 투영한 것이다[6]. 극좌표 형식으로 표현하면 경계부분의 불연속적인 선도 쉽게 표현할 수 있다. 움직이는 물체의 한 화소의 시간적, 공간적 기울기의 변화는 다르지만 움직이는 물체 내에서 같은 방향으로 움직이므로 운동 제한선은 속도 (u′,v′)에서 만난다. 이때 (u′,v′)는 그 화소와 주변 화소의 실제 움직이는 속도이다.

그림 4-2(a)는 밝기 함수가 선형적인 시험 패턴(test pattern)의 이동을 나타내고, (b)는 제한선들의 교차와 d, α를 나타낸다. 밝기 함수가 선형적일 경우 정확한 이동 속도를 구할 수 있다. 그러나 자연 영상은 비선형이므로 교차점이 유일하게 존재하지 않고 흩어져있다. 영상을 가능한 선형적으로 바꾸기 위해 전처리로 가우시안 필터링을 한다.



(a) 선형적인 영상 이동



(b) (a)에 대한 제한선

그림 4-2 선형적인 영상의 이동

Fig. 4-2 Motion of a linear image

제한선의 교점은 극좌표의 좌표축 회전을 통해 쉽게 구할 수 있다. 두 제한선이 그림 4-3와 같다고 하자.

$$b = \frac{c}{\tan \Phi} \tag{4-5}$$

c는 d, d', $\Phi = \alpha' - \alpha$와 다음과 같은 관계가 있다.

$$(d+c)\cos \Phi = d' \tag{4-6}$$

식(4-5)를 식(4-6)에 대입하여 b에 대해서 풀면,

$$b = \frac{d' - d\cos \Phi}{\sin \Phi} = \frac{d' - d\cos(\alpha' - \alpha)}{\sin(\alpha' - \alpha)} \tag{4-7}$$

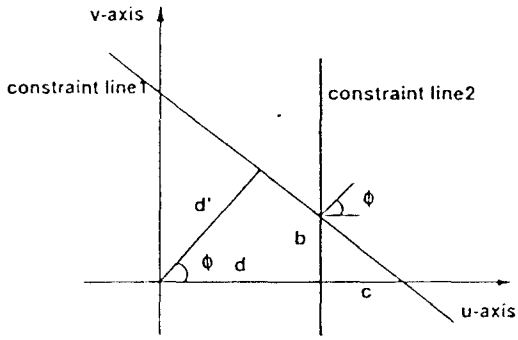


그림 4-3 제한선의 교점

Fig. 4-3 Intersection point of constraint lines

이 된다. 여기서 d와 b는 각각 u,v축의 속도이다. 이와 같이 이웃 제한식들과의 교차점을 구한다. 구해진 이동 속도는 국부적인 값이기 때문에 광역정보가 일부 손실된다. 특히 경계 부분과 모서리에서 심하게 나타난다.

2. 수정된 계층구조의 모델과 알고리즘

제한선의 교차점은 국부적인 부분의 이동 속도이므로 실제로 움직인 물체는 대부분 일정한 방향인데 반하여, 속도 (u,v)는 독립성을 가지므로 불규칙적으로 나타난다. 또한, 부적절한 배치로 인한 영향으로 이동 속도는 기울기 변화에 매우 민감하고 이로 인하여 구경 문제가 발생한다. 이와 같은 문제를 해결하기 위해 광역 전파특성이 좋은 피라미드 구조를 이용하여 반복적 접근식으로 속도를 전체적으로 전파시킨다.

피라미드 구조는 계층이 많을수록 전술한 간축과 대역제한으로 인하여 잡음이 발생하고, 정보의 상실량이 커지게 된다. 자연 영상과 같은 고주파 성분이 많으면 더욱 심하다. 간축으로 인한 오차도 줄이고, 제한식으로부터 구한 속도를 반복적 접근 방법으로 광역으로 완만하게 전파시키기에 알맞은 3층 피라미드 구조를 선택하였다(그림 4-4).

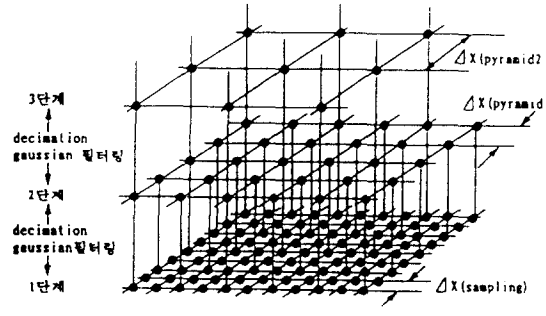


그림 4-4 사용된 피라미드 구조

Fig. 4-4 The pyramid structure used in the algorithm

제한식으로부터 구한 이동 속도 (u,v)를 4×4 크기로 확률밀도 함수 (P_u,P_v)로 나타내어 가장 큰 확률분포값에 해당하는 (u,v)를 구한다. 이 (u,v)를 3층 피라미드 반복접근식의 초기값으로 사용하여 수정 알고리즘에 따라 이동 속도를 추정해 간다.

수정 알고리즘

단계 1. Calculate constraint line

단계 2. Calculate ($u^{(n)}$,$v^{(n)}$) from the constraint lines and calculate the probability density ($P_u^{(n)}$,$P_v^{(n)}$) within each 4×4 window size

단계 3. ($u'^{(n)}$,$v'^{(n)}$) = select (u,v) of max(P_u,P_v) for each windows

단계 4. $u^{proj} = u'^{(n)}$, $v^{proj} = v'^{(n)}$, k=L

단계 5. Calculate

$$P = I_x^{proj} + I_y^{proj} + I_t$$

$$D = \alpha^2 + I_x + I_y$$

$$u^k = u^{proj} - I_x \frac{P}{D}$$

$$v^k = v^{proj} - I_y \frac{P}{D}$$

단계 6. $u^{proj} = P_k^{k-1}(u^k)$, where $P_n^m(\,)$ is a projection n operator $v^{proj} = P_k^{k-1}(v^k)$ of level n into m

k = k - 1

단계 7. Calculate

$$P = I_x^{proj} + I_y^{proj} + I_t$$

$$D = \alpha^2 + I_x + I_y$$

$$u^k = u^{proj} - I_x \frac{P}{D}$$

$$v^k = v^{proj} - I_y \frac{P}{D}$$

단계 8. $u^{proj} = P_k^{k-1}(u^k)$

$v^{proj} = P_k^{k-1}(v^k)$

k = k - 1

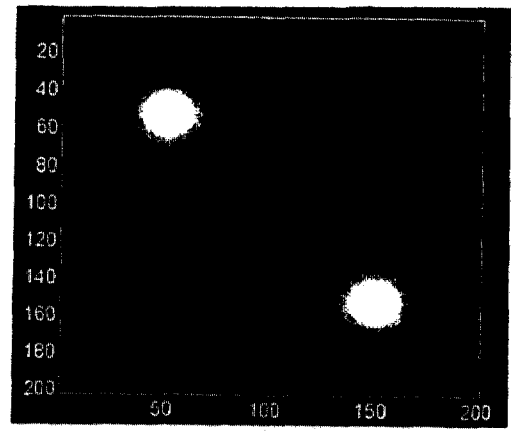
단계 9. If (converge) k = k - 1

If (k = 1) stop

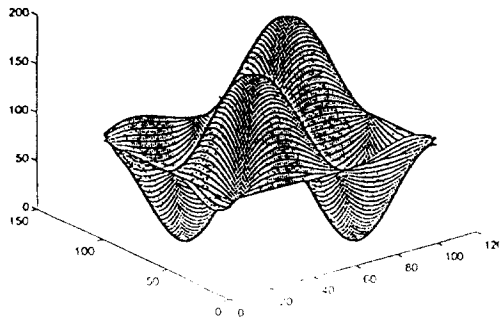
else goto step 2

V. 실험 및 결과 비교 검토

본 논문에서는 제안된 알고리즘의 성능 측정을 위해 인위적인 영상에 대한 컴퓨터 시뮬레이션을 수행하고 이동 변위에서 FD와 같은 기존의 계층 구조의 변위 측정 알고리즘과 비교, 검토한다. 그림 5-1은 인위적으로 생성시킨 시험 패턴으로 정중앙을 중심으로 5° 회전한 움직임 영상이다.

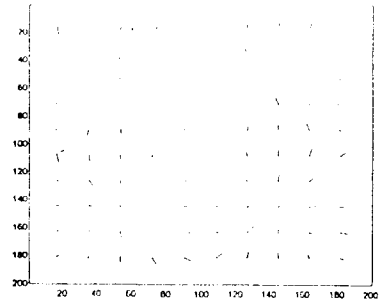


(a) 시험 영상

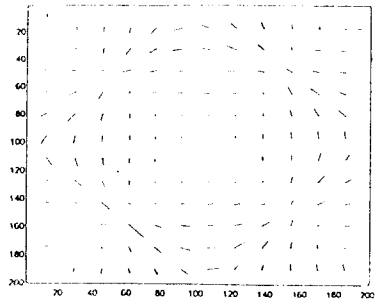


(b) 회전 영상의 삼차원 표현
그림 5-1 시험 영상
Fig. 5-1 test pattern

FD 시뮬레이션 결과(그림 5-2)와 실제 움직인 속도와 비교해 보면 많은 차이가 있다. 이는 회전 영상의 부적절한 배치로 인한 구경 문제가 심각하기 때문이다. 또한 수렴 속도는 공간 경사보다 시간 경사에 더욱 민감하다.



(a) α=0.001, 반복 횟수=16



(b) α=0.001, 반복 횟수=32
그림 5-2 FD 이동 속도 추정
Fig. 5-2 The estimation of FD motion velocity

부적절한 배치로 인한 구경 문제를 해결하기 위해 Ⅲ장에서 논한 계층 구조를 이용하여 이동 속도를 추정하면 그림 5-3과 같다. FD와 결과를 비교해 보면 FD보다 더 좋은 결과를 얻을 수 있음을 알 수 있다. 그러나 가장자리 부분에 오차를 함축하고 있다.

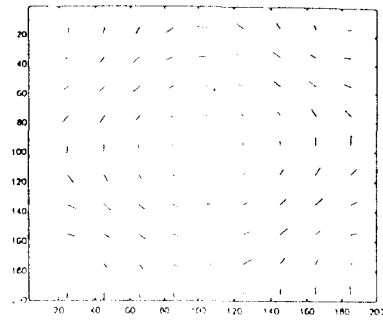


그림 5-3 계층 구조를 이용한 이동 속도 추정(α=0.001)
Fig. 5-3 The estimation of motion velocity using the hierarchial structure

4장에서 제안한 알고리즘에서 사용한 운동 제한선의 교점(이동 속도)을 구해보면 그림 5-4과 같다.

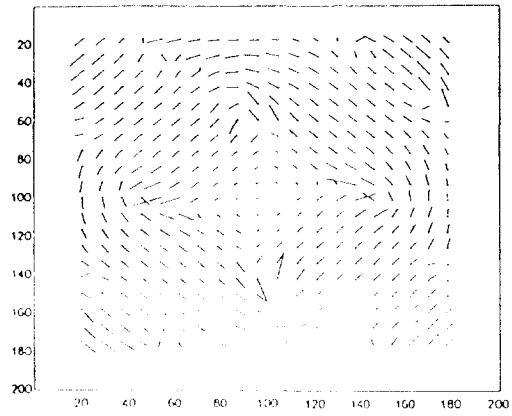


그림 5-4 운동 제한선의 교점에 의한 이동 속도 추정
Fig. 5-4 The estimation of the motion velocity by intersections of con-straint lines

운동 제한선의 교점들을 V.2절에서 제안한 식을 사용하여 구해보면 그림 5-5과 같이 가장자리에서 조금 흐릿해짐(blurring) 현상이 나타나고, 시간 경사의 영점 교차 부분에서 불연속성이 나타난다. 이를 줄이기 위해서는 적당한 반복 횟수와 α를 결정해야 한다. 선형 영상의 움직임에서는 FD, 계층 구조를 이용한 알고리즘의 성능을 비교해 본 결과 정상상태 오차에 일찍 도달하고 정상상태 오차도 적음을 알 수 있다(그림 5-6).

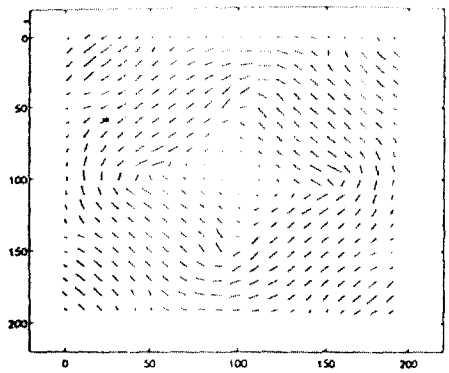


그림 5-5 제안한 알고리즘의 이동 속도 추정 (α =0.001)

Fig. 5-5 Estimation of motion velocity by proposed algorithm

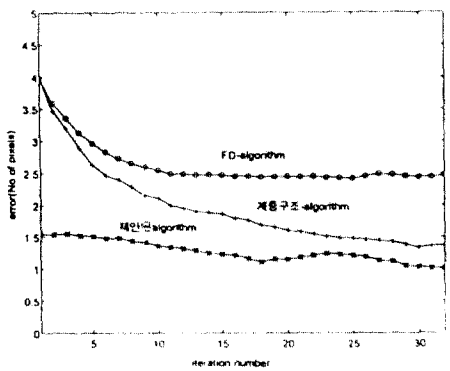


그림 5-6 선형 영상의 정상 상태 오차 (α =0.001, iteration=32)

Fig. 5-6 Steady state error of linear image(α =0.001, iteration=32)

VI. 결 론

본 논문에서는 최근 들어 활발히 연구되고 있는 운동영상 해석방법 중에서 계층구조방법의 개선을 논의하였다. 물체의 밝기를 근거로 하는 방법에서 기본이 되는 공간 경사와 시간 경사의 측정 방법과 경사를 기초로 하는 방법의 측정 시 발생하는 불완전한 조건, 부적절한 배치와 샘플링 오차에 대해 논의하였다. 이들 오차를 개선시키기 위해 인간의 인식 구조와 유사한 계층 구조를 이용한 변위 추정 방법을 사용하였다. 그리고 각 계층별 오차를 해석하였다. 이 계층 구조의 개선책으로 경계 부분에서의 속도의 흐릿해짐을 막기 위해 제한식들의 교점을 이용하여 변위를 구한 다음 이를 반복 접근식의 초기값으로 사용하여 각 계층으로 전파 시킴으로서 광역 정보를 함축시켜 변위 추정 시 수렴 속도를 빨리 하고 정상상태 오차를 줄였다. 그리고 구경 문제로 인한 오차를 줄였다. 다른 알고리즘에 비해 정상상태 오차가 작고 각 계층간 수렴 속도도 빠르다는 결론을 얻었다. 그러나 영상의 불연속적인 부분에서는 변위가 잘못 구해진다. 특히 시간 경사의 영점 교차 부분에서 심하게 나타난다. 이를 개선시키기 위해서는 적절한 가우시안 필터 계수와 반복 접근식의 α를 적절히 결정해야 한다.

앞으로 시간 경사의 영점 교차 오차를 해석하고, 이를 줄이는 알고리즘에 대해 연구가 이루어져야 할 것이다. 그리고 운동 제한식에 어떻게 선형성을 부여할 것인가에 대한 연구가 필요하고, 특징 정합 방법과 기울기를 기초로한 방법을 조합시키는 알고리즘에 대해서도 연구가 되어야 할 것이다. 또한, 카메라를 두 대 이상 이용하는 깊이 정보가 있는 스테레오 영상의 운동 추정에 대해서도 연구가 이루어져야 한다.

<참고문헌>

[1] D. Marr, *Vision*, W.H Freeman and Company, San Fransisco, 1982.

[2] W.D Thomson, K.M. Mutch and U.A. Bertines "Dynamic occlusion analysis in optical flow fields" IEEE trans. on PAMI, Vol. 7, no. 4, pp. 374-383, July, 1985.

[3] G, Adiv, "Determining three-dimensional motion and structure from optical flow generated by several moving objects", IEEE trans. on PAMI, Vol. 7, no. 4, pp. 384-401, 1985.

[4] K. Pradny, "Motion and structure from optical flow" in proc. of Int. Joint Conf. on Artificial Intelligence, pp. 702-704.

[5] A.R. Bruss and B.K.P Horn, "Passive navigation". Computer Vision, Graphics and Image Processing, vol. 21, no. 1, pp. 3-20, Jan. 1987.

[6] K.P. Horn and B.G. Shunk, "Determining optical flow" Artificial Intelligence, vol. 17, pp. 185-204, 1981.

[7] G. Adiv, "Inherent ambiguities in recovering 3-D motion and structure from a noisy flow field" IEEE trans. on PAMI, vol. 11, no. 5, pp. 477-489, 1987.

[8] M.Yachida, "Determining velocity by spatio-temporal neighborhoods from image sequences" Comp. Vision Graphics and Image Processing, vol. 21, pp. 185-203, 1981.

[9] Z. Wu, H. Sun and L.S. Davis, "Determining velocity by propagation" in pro. Int. conf. Pattern Recognition Munich. West Germany pp. 1147-1149, Oct. 19-22 1982.

[10] N. Cornelius and T. Kanade, "Adapting optical flow to measure object motion in reflectance and X-ray image sequences" in proc. ACM siggaph/sigart and perception, Tronto, Ont, Canada, pp. 50-58, Apr. 4-6 1983.

[11] B.G. Shunk "Image flow segmentation and estimation by constrain line clustering" IEEE trans. on PAMI vol. 11, pp. 1010- 1027, Oct. 1989.

[12] A. Spoerri and S. Ullman, "The early detection of motion boundaries" Proc. of International Conference on Computer vision pp. 209-218, Jun. 8-11 1987.

[13] J. Kearney W.B. Thomson, and D.L. Boley, "Optical flow estimation: An error analysis of gradient based methods with local optimization", IEEE trans. on PAMI vol. PA MI-8, no. 5, pp. 565-593, 1986.

[14] H.H. Nagel and W. Enkelman, "An investigation of smoothness constraints for the estimation of displacement vector field from image sequences" IEEE trans. on PAMI, vol. PAMI-8, no. 5, pp. 565-593, 1986.

[15] J.R. Westlake, *A handbook of numerical matrix inversion and solution of linear equation*, New York, 1968.

[16] M.A. Marshall, "Self-organizing neural network architectures for computing visual depth from motion paradox" proc. of Int. Joint Conf. on neural network.

[17] R.N. Bracewell, *The Fourier Transform and its Application*, Mc Graw-Hill, New York, 1978.

[18] P. Marr and E.C. Hildreth, "Theory of edge detection" proc. of R. Soc.

Lond, vol. 207, pp. 178-217, 1980.
[19] P.J. Burt, "Fast filter transforms for image processing" Computer, Graphics and Image processing, vol. 16, pp. 20-51, 1987.