

Optical flow field 분할을 위한 2½D 및 정방형 모델과 결합된 버텀-업 피라미드 링킹 방법에 관한 연구

正會員 金 春 吉* 正會員 李 亨 宰**

A Study on Bottom-up Pyramid Linking (BUPL) Method Combined with 2½D and Quadratic Model for Segmentation of Optical Flow field

Choon Gil KIM*, Hyung Jae LEE** *Regular Members*

要 約 Optical flow는 물체의 속도 결정과 추적 뿐만 아니라 영상분할과 물체의 3차원 정보로서도 중요하다. 본 논문에서는 OFF(optical flow field)의 해석 즉, OFF 분할을 위하여 3가지 모델과 결합된 피라미드 링킹방법의 분할 알고리즘을 제시하였다. 제시된 방법에서 각 노드는 노드가 지배하는 영역내의 flow field의 모델을 형성하고, 이 영역들을 최적으로 결합되는 모델을 취함으로써 상위 레벨상의 노드와 결합되어진다. 또한 각 노드는 노드가 지배하는 화소들의 부모화소의 flow field 모델 사이에서 발생하는 최소제곱오차를 기준으로 하여 부모화소와 링크된다. 이와 같은 방법에 의해 기존 방법들에서의 주된 문제점인 잡음에 대한 민감성을 개선할 수 있었고, 이를 실험을 통하여 입증하였다.

ABSTRACT Optical flow is important not only for determining velocity and trajectory of the object but also for image segmentation and three dimensional information. In this paper, an algorithm for segmentation of the optical flow field is presented. This method is based on a pyramid linking method combined with the proposed models. In this method, each node contained a model of the flow in the region that it represented. Regions were combined by taking the model that best fit the union of the two regions. Each node linked to one of its fathers based on the error between the pixels represented by the node and the father's model of its flow. A major problem which has emerged in conventional researchs on optical flow field is sensitive to noise. The proposed method is relatively insensitive to noise. At the result of computer simulation, the pyramid algorithm proposed in this paper seem to have useful properties.

I. 서 론

일반적으로 정지영상(static scene)을 분할하는 것보다 운동영상(dynamic scene)을 분할하는 것이 영상을 더 의미있는 영역으로 나눌 수 있다. 운동영상에서 물체의 운동(motion) 그 자체 뿐만 아니라 물체의 구조를 알아내는 것은 모든 동물들이 가지고 있는 속성이며 기계(machine)로 하여금 어떤 일(task)을 하도록 할때 줄 수 있는 효율적인 정보이다. 이러한 이유로 해서

근래에 들어 운동영상해석(dynamic scene analysis) 또는 운동해석(motion analysis) 분야에 대한 연구가 급증하고 그의 응용 또한 다방면으로 발전되고 있다. 예를들면, 자동목표 추적장치(automatic object tracker)나 항법장치(navigation), 자동감시장치(automatic surveillance system)와 같이 직접적으로 관련이 있는 분야외에 영상회의장치(video conference), 영상전화와 같은 전송장치와 의학, 무도법(choreography), 기상학 등의 학문적인 분야에 이르기 까지 널리 이용되고 있다¹⁾.

운동영상을 해석하는 방법에 대하여 지금까지 연구된 것을 대체적으로 두가지로 분류할 수 있는데, 특징을 근거로 한 방법(feature based

*韓國科學技術大學

**韓國航空大學

Hankuk Aviation University

論文番號 : 91-109 (1991. 7. 29)

method) 그리고 optical flow에 근거한 방법으로 나눌 수 있다. OFF(Optical flow field)에 관한 연구는 OFF의 계산과 OFF의 해석 즉, 분할로 세분화될 수 있다. OFF 연구에서 OFF는 화소 단위의 순간 속도를 묘사하는 벡터로 구성된 속도장(velocity field)과 과거 영상과 현 영상 사이에서 화소의 변위를 묘사하는 벡터로 구성된 변위장(displacement field)을 의미한다. 본 논문에서는 전술한 OFF(optical flow field)의 해석 즉, OFF의 속도장(velocity field)에 관한 영상 분할에 관하여 논하고자 한다. OFF의 분할은 순차적인 flow field 영상들의 이용과는 반대로 단 하나의 flow field 영상내의 정보만을 이용한다⁽²⁾.

OFF 해석에 관한 기존 연구에서 제기되는 주된 문제점은 잡음에 대한 민감성이다^{(9) (8)}. 기존 기술들에 의해 생성된 flow field는 잡음이 존재하고 부분적으로 왜곡된 상태이다⁽³⁾. 그러므로 이러한 조건하에서 flow field 해석을 위해 제안된 기존 방법들은 많은 문제점을 갖고 있으며 영상분할이 제대로 이루어지지 않았다^{(9) (14)}. 또한 이용될 수 있는 모든 정보를 이용하는 전역적(global) 접근방법은 비교적 해석이 강건하나 최적 기준의 부적절한 선택이 이 방법들의 수행을 제한할 뿐만 아니라 독립적으로 움직이는 물체의 존재는 이러한 전역적 방법을 비효율적으로 만들었다^{(13) (15)}.

OFF를 분할할 때 가장 문제시 되는 점은 다음 두 가지로 요약될 수 있다. 첫째가 OFF에 잡음의 존재이고, 두 번째가 독립적으로 움직이는 물체의 존재이다. 본 논문에서는 이러한 두 가지 문제점을 해결하기 위해 2½D 모델과 일반적인 정방형 모델(general quadratic model)을 적용하여 OFF를 분할하기 위한 BUPL(bottom up pyramid linking) 방법을 제안한다. 2½D 모델을 이용한 OFF 분할의 가능성을 1984년에 Waxman이 제안하였다⁽¹³⁾. 본 논문에서 제안된 BUPL 방법은 병렬 형식의 구조이므로 병렬처리식 컴퓨터에 적용되면 $O(\log N)$ 의 수행 절차가 요구되므로 기존방법들 보다 수행속도가

빠르다.

본 논문에서 2½D 모델 사용시 영상내의 한 화소의 움직임은 회전(rotation) 및 팽창(dilation) 이동과 결합된 병진(translation) 이동에 의해 근사되어진다. 즉, 이것은 영상평면에 수직인 축에 대한 회전과 임의의 병진운동을 갖는 영상평면에 평행인 평면에 대응된다. 이 모델은 망원 렌즈를 통한 영상의 관찰과 같은 시야의 각이 좁을 때 이루어지는 OFF 해석에 적합하다. 정방형 모델에서는 OFF는 평면내에 정방형 함수가 갖는 12개의 파라미터에 의해 표현된다. 만약 영상내의 한 영역이 공간상에서 한 평면을 구성한다면, 이 모델은 완벽하게 적합되어지고 평면의 파라미터와 그의 움직임은 강건(robust)하게 구해진다. 1984년 Waxman은 영상의 윤곽을 구함으로써 이 모델의 파라미터를 계산하였다⁽¹³⁾.

또한 본 논문에 사용된 피라미드 구조에서 한 노드의 값은 그 노드와 링크된 입력 영상내의 모든 화소의 값과 최소제곱오차에 의해 적합이 된다. 각 노드는 노드가 지배하는 화소들과 최소 오차를 갖는 부모와 링크되어지고, 요구되는 모든 계산은 피라미드내의 국부적인 정보만을 이용하여 처리된다. 이와같은 방법에 의해 형성된 피라미드의 최고레벨상의 노드들이 지배하는 하위레벨상의 각 노드들을 라벨링(labeling)함으로써 영상을 분할한다.

II 장에서는 OFF 분할을 위해 사용된 3가지 모델을 제시하고 III 장에서는 피라미드 링크링과정을 서술한다. IV 장에서는 컴퓨터 모의실험을 보이고 V 장에서 결론을 맺는다.

II. OFF(optical flow field) 적합을 위한 모델

1. 2½D 모델

입력영상에 대한 피라미드 구조에 의해 형성된 각 영역에 대한 flow field는 병진(translation), 회전(rotation) 및 팽창(dilation)운동에 의해

해석되어진다. 이 모델에서 임의 한 화소좌표 (x, y)에서의 속도(velocity)는 식(1)과 같다.

$$\begin{aligned} V_x &= T_x - D_x - R_y \\ V_y &= T_y - D_y + R_x \end{aligned} \tag{1}$$

여기서 (T_x, T_y)는 평진운동 벡터이고, 회전과 평창운동은 D와 R에 비례한다. 영상에서 좌표 (x_i, y_i)에 위치한 화소에 대하여 벡터 (O_x, O_y)가 주어진다고 가정하면, 이 좌표에서 관측속도와 계산속도의 근사치와의 차는 식(2)와 같이 주어진다.

$$\begin{aligned} E_i &= (T_x - D_{xi} - R_{yi} - O_{xi})^2 + \\ &(T_y - D_{yi} + R_{xi} - O_{yi})^2 \end{aligned} \tag{2}$$

영상내의 영역에 대한 총 오차 제곱의 합은 식(3)과 같다.

$$\begin{aligned} E^2 &= \sum_{i=1}^k ((T_x - D_{xi} - R_{yi} - O_{xi})^2 + \\ &(T_y - D_{yi} + R_{xi} - O_{yi})^2) \end{aligned} \tag{3}$$

여기서 k는 영역내의 화소의 수이다. 이 오차를 최소화하는 파라미터의 값을 구하기 위하여 식(4)와 같이 각 파라미터에 대한 식(3)의 편미분을 0로 한다.

$$\begin{aligned} \frac{\partial E^2}{\partial T_x} &= 2 \sum_{i=1}^k [T_x - D_{xi} - R_{yi} - O_{xi}] \\ &= 0 \end{aligned}$$

$$\begin{aligned} \frac{\partial E^2}{\partial T_y} &= 2 \sum_{i=1}^k [T_y - D_{yi} + R_{xi} - O_{yi}] \\ &= 0 \end{aligned}$$

$$\begin{aligned} \frac{\partial E^2}{\partial D} &= -2 \sum_{i=1}^k [x_i(T_x - D_{xi} - R_{yi} - O_{xi}) + \\ &y_i(T_y - D_{yi} + R_{xi} - O_{yi})] \\ &= 0 \end{aligned}$$

$$\begin{aligned} \frac{\partial E^2}{\partial T_x} &= -2 \sum_{i=1}^k [y_i(T_x - D_{xi} - R_{yi} - O_{xi}) + \\ &x_i(T_y - D_{yi} + R_{xi} - O_{yi})] \\ &= 0 \end{aligned} \tag{4}$$

식(4)로부터 식(5)와 같은 파라미터 값을 구할 수 있다.

$$\begin{aligned} D &= \frac{(x_0x - x_0x) + (y_0y - y_0y)}{(x^2 - x^2) + (y^2 - y^2)} \\ R &= \frac{(y_0x - y_0x) + (x_0y - x_0y)}{(x^2 - x^2) + (y^2 - y^2)} \end{aligned}$$

$$\begin{aligned} T_x &= O_x + R_x + D_y \\ T_y &= O_y + R_y - D_x \end{aligned} \tag{5}$$

여기서 바(bar) 기호는 영역내에 존재하는 화소들이 갖는 정보의 평균을 의미한다.

예를들면,

$$\bar{x} = \frac{1}{k} \sum_{i=1}^k x_i \quad (6)$$

이다. 식(5)는 영역에 최적 적합(best fit)되는 근사치를 구하기 위해 필요한 정보가 10개의 값 즉, $x, y, x^2, y^2, O_x, O_y, xO_x, xO_y, yO_x$ 및 yO_y 임을 의미한다. 각 노드를 위한 이 값들은 피라미드 형성시 각 노드의 자식화소들을 평균함으로써 국부적인 방법에 의해 구해지고, 두 개의 상의 화소가 하나의 노드를 형성하면 근사치는 정확하게 계산되어진다.

일단 피라미드내의 각 노드에 대한 optical flow가 계산되어지면, 각 노드는 관련된 4개의 부모화소와 비교되어지고, 이 때 발생하는 오차는 부모화소의 값 T_x, T_y, R, D 를 이용하여 식(3)의 합을 전개하므로써 구해질 수 있다.

2. 정방형(quadratic) 모델

2½D 모델 대신에 일반적인 정방형 모델을 사용하면, 좌표(x, y)에서의 optical flow에 대한 모델은 식(7)과 같다.

$$\begin{aligned} V_x &= A_x + B_x x + C_x y + D_x x^2 + E_x xy + F_x y^2 \\ V_y &= A_y + B_y x + C_y y + D_y x^2 + \\ &E_y xy + F_y y^2 \end{aligned} \quad (7)$$

여기서 12개의 값 A_x, C_x, D_x, E_x, F_x 및 $A_y, B_y, C_y, D_y, E_y, F_y$ 는 모델의 파라미터이다. 근사치의 총 오차 제곱의 합(total squared error)

$$E^2 = \sum_{i=1}^k (A_x + B_x x_i + C_x y_i + D_x x_i^2 +$$

$$\begin{aligned} &E_x x_i y_i + F_x y_i^2 - O_x i)^2 \\ &+ \sum_{i=1}^k (A_y + B_y x_i + C_y y_i + D_y x_i^2 + E_y x_i y_i + \\ &F_y y_i^2 - O_y i)^2 \end{aligned} \quad (8)$$

은 식(8)과 같이 주어진다.

여기서 k는 영역내의 화소들의 수이고, 식(8)은 6개씩의 파라미터로 구성된 2개의 독립적 부분으로 구성되어 있으므로 파라미터 $A_x, B_x, C_x, D_x, E_x, F_x$ 를 구하는 방법은 나머지 6개의 파라미터를 구하는 방법과 동일하다. 그러므로 본 논문에서는 x 에 관한 파라미터들만 다룬다. 데이터에 최적 적합되는 파라미터를 구하기 위하여 2½D 모델에서와 같이 각 파라미터에 대한 식(8)의 편미분을 0으로 하면 식(9)와 같이 된다.

$$\begin{aligned} \frac{\partial E^2}{\partial A_x} &= \sum_{i=1}^k 2(A_x + B_x x_i + C_x y_i + D_x x_i^2 + \\ &E_x x_i y_i + F_x y_i^2 - O_x i) \\ &= 0 \\ \frac{\partial E^2}{\partial B_x} &= \sum_{i=1}^k 2x_i (A_x + B_x x_i + C_x y_i + D_x x_i^2 + \\ &E_x x_i y_i + F_x y_i^2 - O_x i) \\ &= 0 \\ \frac{\partial E^2}{\partial C_x} &= \sum_{i=1}^k 2y_i (A_x + B_x x_i + C_x y_i + D_x x_i^2 + \end{aligned}$$

$$E_{xx_i y_i} + F_{xy_i^2} - O_{x_i}) \\ = 0$$

$$\frac{\partial E^2}{\partial D_x} = \sum_{i=1}^k 2x_i^2 (A_x + B_{xx_i} + C_{xy_i} + D_{xx_i^2} +$$

$$E_{xx_i y_i} + F_{xy_i^2} - O_{x_i}) \\ = 0$$

$$\frac{\partial E^2}{\partial E_x} = \sum_{i=1}^k 2x_i y_i (A_x + B_{xx_i} + C_{xy_i} +$$

$$D_{xx_i^2} + E_{xx_i y_i} + F_{xy_i^2} - O_{x_i}) \\ = 0$$

$$\frac{\partial E^2}{\partial F_x} = \sum_{i=1}^k 2y_i (A_x + B_{xx_i} + C_{xy_i} + D_{xx_i^2} +$$

$$E_{xx_i y_i} + F_{xy_i^2} - O_{x_i}) \\ = 0 \tag{9}$$

식(9)를 정리하면 식(10)이 주어진다.

$$A_x + B_{xx} + C_{xy} + D_{xx^2} + E_{xxy} + \\ F_{xy^2} - O_x = 0$$

$$A_{xx} + B_{xx^2} + C_{xxy} + D_{xx^3} + E_{xx^2y} + \\ F_{xxy^2} - xO_x = 0$$

$$A_x y + B_{xxy} + C_{xy^2} + D_{xx^2y} + E_{xxy^2} + \\ F_{xy^3} - yO_x = 0$$

$$A_{xx^2} + B_{xx^3} + C_{xx^2y} + D_{xx^4} + E_{xx^3y} + \\ F_{xx^2y^2} - x^2O_x = 0$$

$$A_{xxy} + B_{xx^2y} + C_{xxy^2} + D_{xx^3y} + E_{xx^2y^2} + \\ F_{xxy^3} - xyO_x = 0$$

$$A_{xy^2} + B_{xxy^2} + C_{xy^3} + D_{xx^2y^2} + E_{xxy^3} + \\ F_{xy^4} - y^2O_x = 0 \tag{10}$$

식(10)은 피라미드 형성시 화소ⁱ에 대한 계산된 값의 평균만을 이용하므로 파라미터들의 값은 쉽게 구해진다. 2½D 모델의 경우와 같이 부모화소와 자식화소 사이의 오차는 식(8)을 전개하므로써 구해진다.

3. 한 성분이 주어진 경우의 2½D 모델

임의 위치에서 속도에 대한 한 성분이 주어질 경우 단위 벡터(unit vector)와 속도를 dot product 하면 모델에서 발생하는 오차를 구할 수 있다. 즉, 화소 i에서의 이 값과 모델에서 예측되는 값 사이의 오차는 식(11)과 같다.

$$E = V_{model} \cdot U_i - m_i \\ = (T_x - D_{xi} - R_{yi})U_{i,x} + \\ (T_y - D_{yi} + R_{xi})U_{i,y} - m_i \tag{11}$$

여기서 U_i 는 화소 i에서 주어진 속도 방향에서의 단위 벡터이고, m_i 는 영상에서 주어진 성분의 크기이다. 파라미터에 의해 형성된 영역에 대한

오차 제곱의 합은 식(12)와 같다.

$$E^2 = \sum_{i=1}^k ((T_x - Dx_i - Ry_i)U_{i,x} + (T_y - Dy_i + Rx_i)U_{i,y} - m_i)^2 \quad (12)$$

여기에서 k는 영역내의 화소의 수이다. 이 오차를 최소화하는 파라미터들의 값을 구하기 위하여 식(13)과 같이 각 파라미터에 대한 식(12)의 편미분을 0으로 한다.

$$\frac{\partial E^2}{\partial T_x} = 2 \sum_{i=1}^k U_{i,x} ((T_x - Dx_i - Ry_i)U_{i,x} + (T_y - Dy_i + Rx_i)U_{i,y} - m_i) = 0$$

$$\frac{\partial E^2}{\partial T_y} = 2 \sum_{i=1}^k U_{i,y} ((T_x - Dx_i - Ry_i)U_{i,x} + (T_y - Dy_i + Rx_i)U_{i,y} - m_i) = 0$$

$$\frac{\partial E^2}{\partial D} = -2 \sum_{i=1}^k (x_i U_{i,x} + y_i U_{i,y}) ((T_x - Dx_i - Ry_i)U_{i,x} + (T_y - Dy_i + Rx_i)U_{i,y} - m_i) = 0$$

$$\frac{\partial E^2}{\partial R} = -2 \sum_{i=1}^k (y_i U_{i,x} - x_i U_{i,y}) ((T_x - Dx_i - Ry_i)U_{i,x} + (T_y - Dy_i + Rx_i)U_{i,y} - m_i) = 0 \quad (13)$$

식(13)을 정리하면 식(14)와 같이 된다.

$$\begin{aligned} U_x^2 T_x + U_x U_y T_y - (x U_x^2 + y U_x U_y) D + (x U_x U_y - y U_x^2) R - m U_x &= 0 \\ U_x U_y T_x + U_y^2 T_y - (y U_y^2 + x U_x U_y) D + (x U_y^2 - y U_x U_y) R - m U_y &= 0 \\ -(x U_x^2 + y U_x U_y) T_x - (y U_y^2 + x U_x U_y) T_y + (x^2 U_x^2 + 2xy U_x U_y + y^2 U_y^2) D - (x^2 U_x U_y - xy U_x^2 + xy U_y^2 + y^2 U_x U_y) R + (mx U_x + my U_y) &= 0 \\ (x U_x U_y - y U_x^2) T_x + (x U_y^2 - y U_x U_y) T_y - (x^2 U_x U_y - xy U_x^2 + xy U_y^2 + y^2 U_x U_y) D + (x^2 U_y^2 - 2xy U_x U_y + y^2 U_x^2) R - (mx U_y - my U_x) &= 0 \end{aligned} \quad (14)$$

식(14)는 4개의 미지수를 갖는 4개의 식으로 구성되어 있기 때문에 파라미터들의 값은 쉽게 구해 질 수 있다. 또한 이 식은 입력영상으로부터 계산되어지는 값들의 평균인 24개의 양으로 표현되어 있다. 이 경우에도 2½D와 정방형 모델의 경우와 같이 부모화소와 자식화소 사이에서 발생하는 오차는 식(12)를 전개하므로써 구해진다

다.

III. OFF 분할을 위한 BUPL 영상의 형성

본 절에서는 입력영상에 대한 지해상도 영상들로 구성된 피라미드 구조를 형성하는 방법과 영상내의 OFF를 분할하는 방법을 제시한다.

그림 1과 같이 지수적으로 감소하는 노드의 정방형태의 배열더미를 피라미드라 하며 이때 각 배열들을 피라미드의 레벨이라 한다¹⁰⁾.

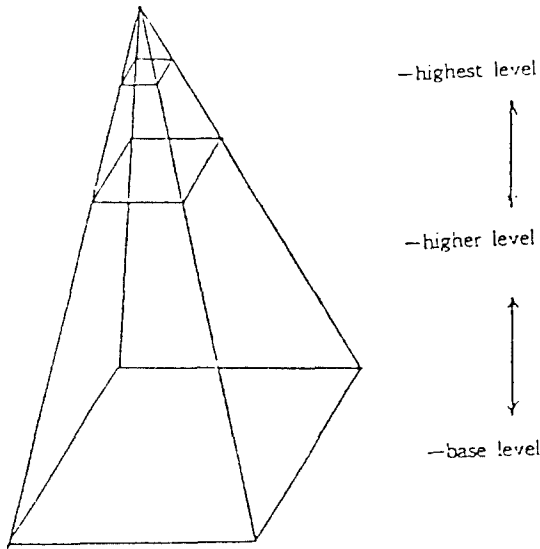


그림 1. 피라미드 영상의 구조
Fig. 1. The structure of pyramid image.

가장 큰 배열을 0레벨 즉, 기저레벨이라 하고, 각 노드는 그림 2와 같이 그 자신의 레벨에서 그의 이웃(north, south, east, west)과 연결되어 있다. 기저레벨위에 존재하는 노드 N은 그의 무레벨 노드들의 클릭과 연결되어지고, 이 때 이 노드들을 N의 자식(children)이라 하며 N은 그들의 부모(parents)라 한다. 해석을 간소화하기 위하여 모든 클릭은 레벨의 가장자리를 제외하고는 같은 형태와 크기를 갖는다고 가정하면 모든 노드는 같은 수 C개의 자식을 갖는다. 클릭이 중첩된다면 한 노드는 하나 이상의 부모를 가질 수 있으며 최고레벨에서의 노드만 제외하고 모든 노드는 적어도 하나 이상의 부모를 갖는다. 다시말해, 모든 노드가 같은 수인 P개의 부모를 갖는다고 가정하면 지수적으로 감소하는 것을 보장하기 위하여 C는 P의 곱 형태가 되어야 한다. 즉 $K > 1$ 인 경우 $C = KP$ 이다. 결국 노드의 수는 레벨과 레벨 사이에서 K의 율로 감소할 것이다.

피라미드의 형성 예로서 레벨들이 $2^n \times 2^n$, $2^{n-1} \times 2^{n-1}$, ...의 크기를 갖고 레벨 k에서 노드 (i, j)의 자식들이 레벨 k-1에서의 화소 SW (2i, 2j), SE(2i+1, 2j), NW(2i, 2j+1), NE(2i+1, 2j+1)이 되면, 각 노드는 단지 하나의 부모를 갖고 $K=4$ 이며 자식의 클릭이 비중첩되므로 이러한 방식으로 형성된 피라미드를 표준 비중첩형 피라미드(standard nonoverlapped pyramid)라 한다. 또한 기저레벨 위에 존재하는 각 노드가 4×4 클릭의 자식화소를 지배하는 경우 이

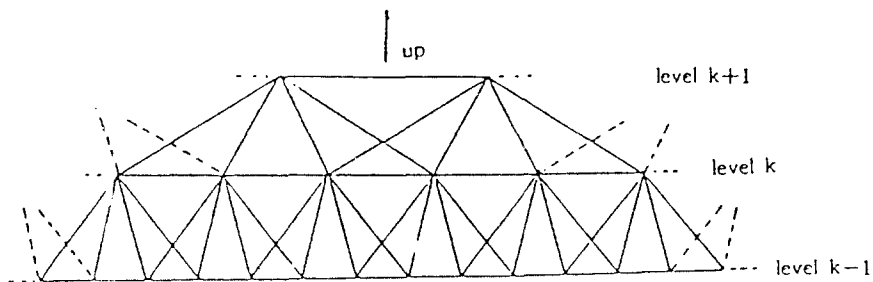


그림 2. 축소되는 레벨사이에서의 1차원적 관계
Fig. 2. One dimensional relational between reduced levels.

블럭은 좌표 i, j 방향 모두에서 50% 중첩현상이 발생된다. 이것은 각 노드가 4부모와 연결됨을 의미하며 $K=4$ 가 됨을 알 수 있다. 이러한 경우 형성된 피라미드를 표준 중첩형 피라미드(standard overlapped pyramid)라 한다⁽¹⁷⁾.

결국 피라미드 데이터구조를 이용하여 영상을 처리할 때 입력된 영상은 노드당 한 화소로 처리되는 피라미드의 기저레벨이 되고, 영상처리는 연속적인 동기식 단계로 이루어진다고 생각할 수 있으며, 각 단계에서 모든 노드는 그의 자식 화소, 이웃 화소, 부모화소로 부터 정보를 얻고, 얻어진 정보에 평균, 최대값 등의 통계적 방법을 이용하여 계산과정을 수행한다.

본 논문에서는 식(15)에 의해 하위레벨에서 상위레벨로 영상을 축소할 때 $2^n \times 2^n$ 의 크기를 가지는 영상은 $2^{n-1} \times 2^{n-1}$ 의 크기로 축소된다.

$$I_k(x, y) = \frac{1}{4} \sum_{i=0}^3 \sum_{j=0}^3 I_{k-1}(x+i, y+j) \quad (15)$$

여기서 $I_k(x, y)$ 는 k 레벨에 있는 좌표 (x, y) 에서의 노드가 갖는 정보를 표시한다. 즉, 영상을 축소할 때 그림 2와 같이 k 레벨의 한 노드는 그의 $k-1$ 레벨의 16화소를 지배하고, 또한 $k+1$ 레벨의 한 노드에 그 영향을 미치는 표준 중첩방법과 각 노드가 갖고 있는 정보를 평균방식을 이용하여 계산한다. 이 때 각 노드는 노드가 지배하는 영역내의 움직임의 근사치를 구하기 위해 필요로 되는 정보를 포함한다. 이 정보는 입력영상의 각 화소에서 계산되어질 수 있는 정보의 평균으로 구성된다. 필요로 되는 정보는 입력영상에서 이용될 수 있는 정보와 II절에서 제시된 OFF에 적용되는 흐름의 모델에 의존한다.

일단 피라미드내의 각 노드가 그의 자식화소가 갖고 있는 정보를 평균하여 초기화되어지면 어떤 링크도 변화되지 않을 때 까지 다음 3단계가 반복된다. 각 노드의 OFF가 노드가 갖고 있는

정보로 부터 근사되어진다. 각 노드는 그의 4개의 부모중의 한 부모와 링크되어진다. 이때 선택된 부모는 노드에 의해 표현되는 OFF와 부모의 근사치와 최적 적합되는 부모가 된다. 만약 부모의 근사치가 결정되지 않는다면 이 부모는 고려 대상에서 제거된다. 각 노드는 노드와 링크된 자식의 정보를 평균함으로써 재 계산된 정보를 갖게된다.

반복과정이 수렴되어지면, 노드를 연결한 링크는 기저레벨의 노드를 일으키고 최고레벨의 노드를 뿌리(root)로 하는 나무(tree)를 형성한다. 영상의 분할은 최고 레벨상의 4개의 노드가 형성하고 있는 링크에 따라서 각 화소를 라벨링(labeling)하므로써 얻어진다. 각 위치에서 흐름의 근사치는 각 위치에 속해 있는 영역의 근사 흐름의 값을 계산함으로써 구해진다. 만약 4개의 영역 이상이 요구되면, 최고레벨보다 하위레벨에 있는 노드들을 나무의 뿌리로 간주하면 된다. 이와 같이 주어진 영상에 대한 적절한 영역의 수를 자동적으로 계산하는 알고리즘이 Hong⁽¹⁸⁾에 의해 연구되었고 여기서는 이 알고리즘을 언급하지 않는다.

IV. 컴퓨터의 모의 실험

본 논문에서 실험대상으로 사용한 영상은 회전(rotation)하는 배경에 팽창하는 모양의 물체를 내포한 8비트 64×64 크기의 그림 3과 같은 인공영상(synthetic image)과, 이 영상에 각각 25%와 50%의 가우시안 잡음(Gaussian noise)이 첨가된 그림 4와 6 영상이다. 또한 그림 8과 같이 임의 움직임을 갖는 2개의 평면에 대한 OFF를 갖는 영상을 이용하였다.

그림 5는 본 논문에서 제시한 2½D 모델 방법에 의해 그림 4의 잡음 영상을 처리한 결과 영상으로 4096개의 화소에 대해 단지 13개의 화소만이 잘못 분할되었다. 그림 7은 그림 6 영상에 대한 결과 영상이다. 본 논문에서 제시한 분할 방법은 기존 방법의 잡음에 대한 영향을 거의 받지 않음을 위의 실험결과 알 수 있다. 또한

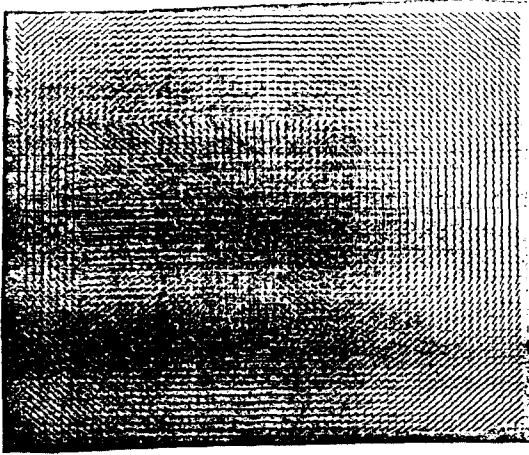


그림 3. 회전모양의 배경에 팽창하는 모양의 OFP를 갖는 인공영상
Fig. 3. Synthetic image with optical flow field showing a dilative object on a rotating background.

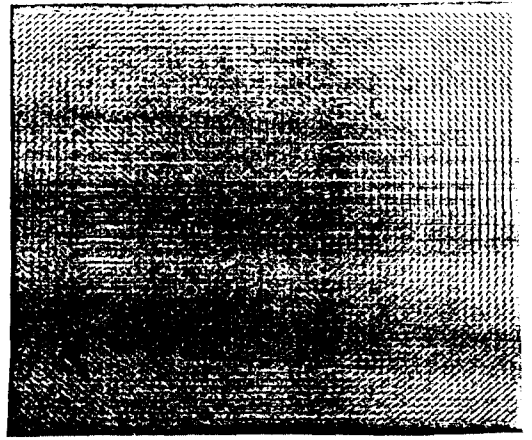


그림 5. 그림 3의 OFP를 분할한 결과영상
Fig. 5. Result image of segmenting Fig. 3.

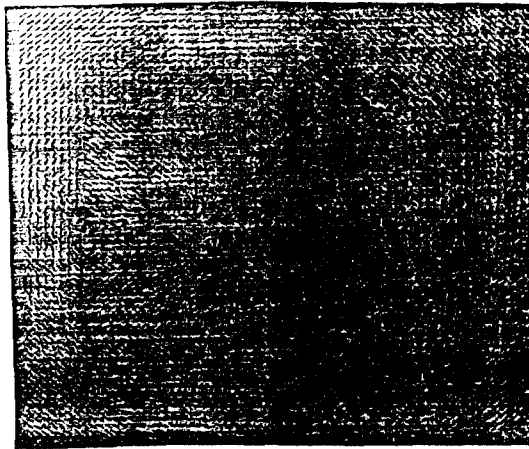


그림 4. 그림 3에 25% 가우시안 잡음이 첨가된 영상
Fig. 4. The image of Fig.3 with 25% Gaussian noise.

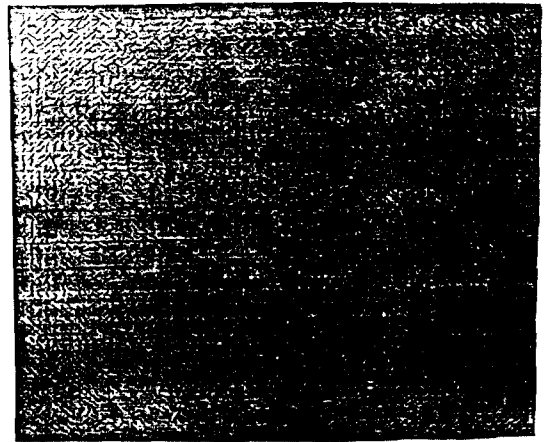


그림 6. 그림 3에 50% 가우시안 잡음이 첨가된 영상
Fig. 6. The image of Fig.3 with 50% Gaussian noise.

그림 8영상에 각각 25%와 50%의 가우시안 잡음이 첨가된 그림 9, 11 영상에 대해 정방향 모델을 적용하여 얻어진 결과영상이 그림 10과 12에 주어졌다. 이 결과영상에서도 위의 모델과 마찬가지로 정방향모델이 잡음의 존재에 강건한(robust)을 알 수 있다.

시간에 대해 순차적으로 얻어진 영상이 주어지면 영상의 한 위치에서 회로의 경도방향(grad-

ient direction)으로 속도 성분은 식(16)을 이용하여 직간에 대한 회로의 미분으로 부터 계산되어 갈 수 있다¹⁰⁾.

$$\frac{aI}{at} = -V \cdot I \quad (16)$$

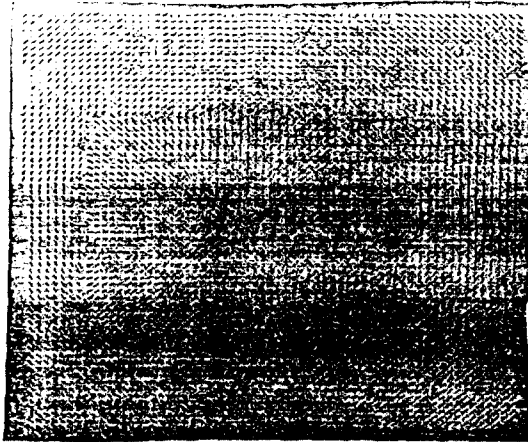


그림 7. 그림 6의 OFF를 분할한 결과영상
Fig. 7. Result image of segmenting Fig. 6.



그림 9. 그림 8에 25% 가우시안 잡음이 첨가된 영상
Fig. 9. The image of Fig.8 with 25% Gaussian noise.

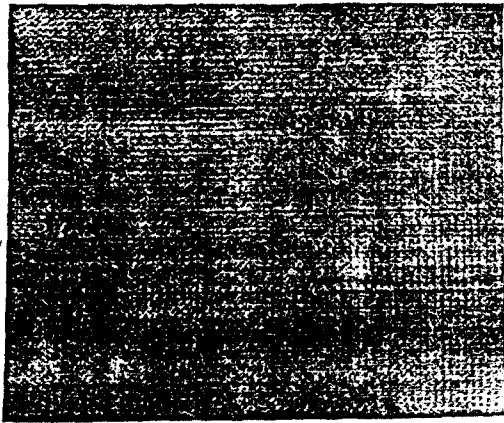


그림 8. 임의 움직임을 갖는 2개 평면에 대한 OFF 영상
Fig. 8. OFF image generated by two moving planes.

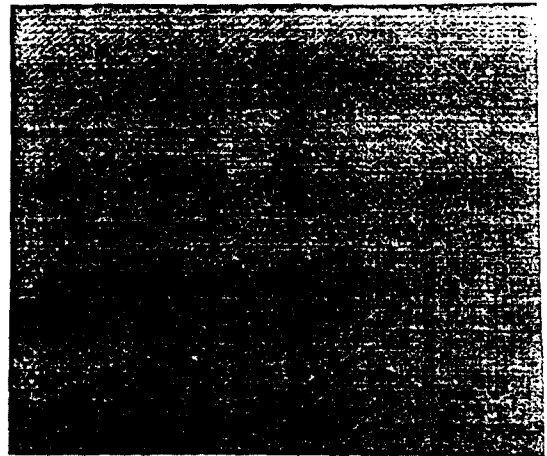


그림 10. 그림9의 OFF를 분할한 결과 영상
Fig. 10. Result image of segmenting Fig.9.

여기서 I 는 휘도값이고, V 는 속도이며 t 는 시간을 의미한다. 그림 13은 정재파 패턴(sinusoidal pattern)과 외부가 약간씩 회전하고 내부는 팽창하는 형태를 갖는 같은 패턴의 영상에서 식(16)으로 부터 유도된 normal 성분을 보여준다. 각 위치에서 시간에 대한 휘도의 미분은 두 프레임 사이에서 휘도차로서 계산되어지고, 경도는 두 프레임사이의 평균(일반적인 3×3 연산자)에 의해 계산되어진다. 본 논문에서 제시한 각 위치에서 한 성분만 주어진 경우의 모델을

사용하여 그림 13을 처리하면 그림14와 같은 영상을 얻을 수 있다. 결과 영상에서 분할이 정확히 발생하지는 않았지만 일반적인 형태는 구분할 수 있었다.

영상이 N 개의 화소를 가지면 피라미드는 $4(N-1)/3$ 개의 노드를 가지기 때문에 일반적인 직렬형 컴퓨터에서 본 논문에 사용된 알고리즘은 영상의 크기에 비례하는 메모리를 필요로하고, 또한 영상의 크기에 비례하는 처리시간을 요구한다. 그러나 각 노드를 분리 처리하는 병렬식

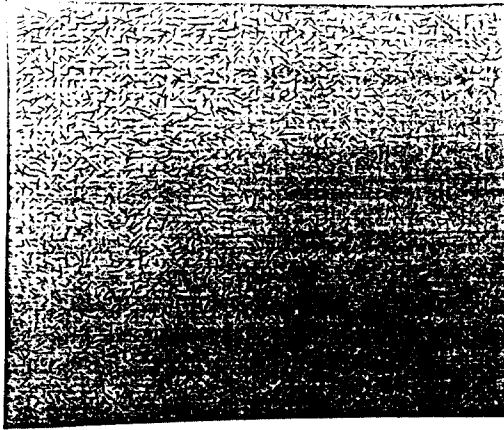


그림 11. 그림 8에 50% 가우시안 잡음이 첨가된 영상
Fig. 11. The image of Fig. 8 with 50% Gaussian noise.

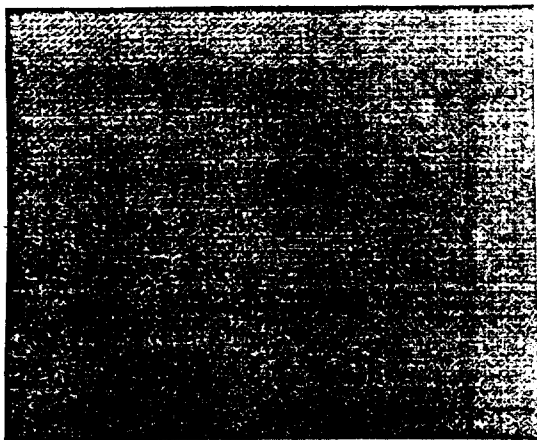


그림 12. 그림 11의 OFF를 분할한 결과 영상
Fig. 12. Result image of segmenting Fig.11.

컴퓨터를 이용하면 임의 레벨에서 모든 노드들은 한 번에 처리됨으로 각 레벨의 처리시간은 $\log(N)$ 에 비례하고, 피라미드의 최고 레벨로 정보가 전달되기 위하여 필요한 시간까지 합하면 총 처리시간은 $O(\log(N))$ 이 요구된다.

본 논문에서 사용된 피라미드 링킹방법은 Iso data 알고리즘의 특별한 경우로 간주 될 수 있다. 이것은 한 위치에서 flow field를 결정하기에 충분한 정보를 갖지 못하는 flow field의 경우에 특히 중요하다. 즉, 영역이 콤팩트하게 된다는

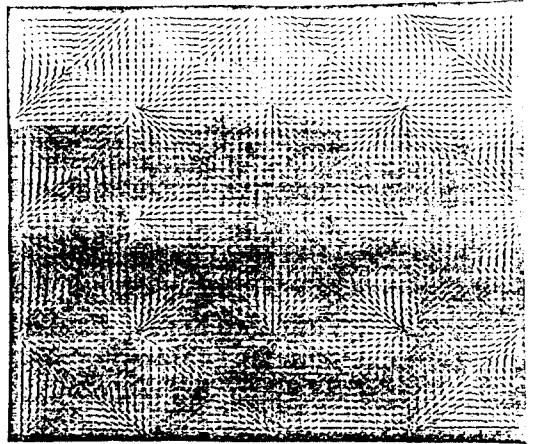


그림 13. 외부가 임의의 회전하고 대부분은 팽창하는 모양을 갖는 정제화로 부터 계산된 속도 방향의 속도 성분
Fig. 13. Velocity component along the gradient computed from a sinusoid with the center expanding and the edge rotating.

사실은 잡음과 모호성의 영향을 감소시킨다.

V. 결 론

본 논문에서는 근래에 들어 활발히 연구되고 있는 운동영상해석 방법중 하나인 optical flow에 근거한 optical flow field의 분할에 대하여 연구하였다. 본 논문에서 제시한 방법은 카메라 센서의 움직임에 의해 야기되는 optical flow

field의 해석을 위한 새로운 방법이다. flow field를 독립적으로 움직이는 물체의 집합으로 재구성하기 위하여 flow field를 여러 레벨로 구성하게 하는 피라미드 링킹방법과 각 영역에 최적 적용되는 flow field를 구하기 위하여 3가지 모델을 이용하였다. 해석 과정에서, 각 레벨상의 노드들은 적절한 피라미터 값에 일치하는 다음 레벨상의 노드들과 결합되고, 각 레벨에서 피라미터 값의 계산은 최소 제곱오차에 기초하여 유도된 방정식들의 해를 구함으로써 이루어진다. 제시된 여러가지 방법을 결합하여 효율적이고 효과적인 알고리즘을 개발하였고, 제시된 알고리즘이 삽입이 내포된 flow field를 분할 가능하게 하였음을 실험을 통하여 입증하였다.

참 고 문 헌

1. J.K. Aggarwal and N. Nandhakumar, "On the computation of motion from sequences of images A review", Proc. of IEEE, vol. 76, no. 8, pp. 917-935, Aug. 1988.
2. G. Adiv, "Determining three dimensional motion and structure from optical flow fields", IEEE trans. on PAMI-7, vol. 7, no. 4, pp. 381-401, 1985.
3. S. Ullman, "Analysis of visual motion by biological and computer systems", Computer, vol. 14, pp. 57-69, Aug. 1981.
4. J.Q. Fang and T.S. Huang, "Solving three dimensional small rotation motion equations", Proc. IEEE Conf. Comput. Vis. Pattern Recog., Washington, DC, pp. 253-258, 1983.
5. J.Q. Fang and T. S. Huang, "Estimating 3 D movement of a rigid object:Experimental results", Proc. Int. Conf. Artif. Intell., Karlsruhe, West Germany, pp. 1035-1037, 1983.
6. K. Prazdny, "Egomotion and relative depth map from optical flow", Biol. Cybern., Vol.36 pp.87-102, 1980.
7. J.W. Roach and J.K. Aggarwal, "Determining the movement of objects from a sequence of images", IEEE Trans. Pattern Anal. Machine Intell., vol. PAMI-2, pp. 554-562, Nov. 1980.
8. R.Y. Tasi and T.S. Huang, "Uniqueness and estimation of threedimensional motion parameters of rigid objects with curved surfaces", IEEE Trans. Pattern Anal. Machine Intell., vol. PAMI-6, pp. 13-27, 1984
9. D.T. Lawton, "Motion analysis via local translational processing", Proc. Workshop cmput. Vis., Rindge, NH, pp. 59-72, 1982.
10. D.T. Lawton, "Processing dynamic image sequences from a moving sensor", Ph.D. dissertation, Dep. Comput. Inform. Sci., Univ. Mass., Amherst, 1984
11. K. Prazdny, "Determining the instantancous direction of motion from optical flow generated by a curvilinearly moving observer", Proc. IEEE Conf. Pattern Recog. Image Proc., Dallas, TX, pp. 109-114, 1981.
12. J.H. Rieger and D.T. Lawton, "Determining the instantaneous axis of translation from optical flow generated by arbitrary sensor motion", Proc. Workshop Motion:Representation perception, Toronto, Canada, pp. 33-41, 1983.
13. A.M. Waxman and K. Wahn, "Contour evolution, neighborhood deformation and global image flow: Planar surfaces in motion", Center Automation Res., Univ. Maryland, CAR TR 58, 1984.
14. H.H. Nagel, "On the derivation of 3 D reconstruction from two perspective views", Proc. Int. Joint Conf. Artif. Intell., Vancouver, Canada, pp. 661-663, 1981.
15. A.M. Waxman, "An image flow paradigm", Proc. of the Second IEEE Workshop Comput. Vision: Representation and Control, pp. 49-57, 1984.
16. A. Rosenfeld(ED.), Multiresolution image processing and Analysis, Springer, Berlin, 1984.
17. T.H. Antonisse, "Image segmentation in pyramid", Comput. Vis. Graph. Image Process., vol. CGIP-19, pp. 367-383, 1982.
18. T. Hong, Pyramid methods in image analysis, Ph. D. thesis, Univ. Maryland, 1982.
19. B. Horn and B. Schunck, "Determining optical flow", Artificial Intell., vol. 17, pp. 185-204, 1981.

金 春 吉(Choon Gil KIM) 正會員
韓國科學技術大學教授

李 亨 宰(Hyung Jae LEE) 正會員
1933年 3月 20日生
한국항공대학 전자공학과 졸업.
상희대학교 대학원(공학박사)
한국항공대학 교수