

확산망을 이용한 스테레오 정합에 관한 연구

(A Study on the Stereo Matching Using Diffusion Networks)

李相璨*, 南基坤**, 金在昌**, 姜昌淳***, 鄭斗泳****, 李尙昱*****

(Sang-Chan Lee, Ki-Gon Nam, Jae-Chang Kim, Chang-Soon Kang, Du-Yeong Jung, and Sang-Wook Lee)

요 약

3차원 거리 정보를 획득하기 위한 스테레오 정합에서 가장 중요한 관점은 어떤 정합 척도(matching measure, matching cost)를 가지고 정합시킬 영역의 크기를 얼마로 할 것인가 하는 것이다. 즉, 비교할 두 개의 영상영역의 크기를 어떻게 최적화 할 것인가 하는 문제이다. 본 논문에서는 이러한 문제를 확산망으로 해결하고자 하였다. 확산망은 국부적인 영상영역에서 간단한 처리과정으로 구성되며, 이러한 처리과정을 반복적으로 처리함으로써 정합척도를 모으는 범위를 변화시켜 영역의 크기를 쉽게 조절할 수 있다. 이것은 스테레오 비전시스템이 요구하는 효율성(efficiency) 병렬성(parallelizability) 및 적응성(adaptability) 등이 간단히 실현됨을 보여준다. 본 논문에서는 제시된 여러 확산망의 특징들을 살펴보고, 확산망을 통해 구성된 스테레오 비전 시스템에 합성 영상 및 실제 영상을 적용하였고, 그 결과를 검토하였으며 스테레오 정합에 널리 사용되는 고정된 영상 영역의 크기를 가지고 정합하는 방법과도 비교하여 우수한 성능임을 보였다.

Abstract

One of the central problems in stereo matching is the selection of the optimal window sizes for comparing image regions. The window size must be large enough to include enough variation for reliable matching, but small enough to avoid the effect of projection distortion. This paper discusses these problems with some novel algorithm based on iterative diffusion process at different disparity hypotheses. Also this paper proposes four kinds of diffusion algorithms to preserve discontinuity in stereo matching. We present and discuss extensive empirical results of algorithms based on various sets of synthetic and real images.

* 正會員, 東義工業專門大學 映像情報處理科
(Dept. of Image Processing, Dongeui technical junior college)

** 正會員, 釜山大學校 電子工學科
(Dept. of Elec. Eng., Pusan National Univ.)

*** 正會員, 東義工業專門大學 電子科
(Dept. of Elec. Eng., Dongeui tech. junior college)

**** 正會員, 巨濟專門大學 電子科
(Dept. of Electronics Eng., Geojae junior college)

***** 正會員, 慶尙大學校 情報通信工學科
(Dept. of Information & Communication Eng., Gyeongsang Nat'l Univ.)

接受日字: 1997年9月27日, 수정완료일: 1998年2月3日

I. 서 론

스테레오 비전은 3차원 공간상에 설치된 카메라의 기하학적 특성이 주어진 상태에서 카메라로부터 얻어진 한 장면에 대한 좌,우 두 개의 영상에서 상호간의 정합점을 찾아내고 정합점으로부터 3차원 깊이 정보를 추출하는 일련의 과정을 거친다. 이 때 좌,우 두 영상으로부터 대응점을 추출하는 정합과정이 가장 어렵고, 연산에 시간이 많이 소요된다. 일반적으로 스테레오 정합의 결과는 영상에서 얻어진 정합점의 상대적 거리를 변이도(disparity map)로 나타낸다.^[1]

스테레오 정합을 위한 방법으로는 크게 특징기반 방법(feature based approach)과 영역기반 방법(area based approach)이 있다. 특징기반 방식에서는 먼저 영상으로부터 안정된 영상특징을 추출하고 추출된 영상특징 공간에서 정합시키는 것으로 처리속도를 향상시킬 수 있으나 영상의 전체 영역에 대한 변이 추정 불가능하며, 전체 영상에 대한 조밀한 변이도를 구하기 위해서는 내삽과정(interpolation)을 거쳐야 하며, 이러한 과정은 영상의 밝기정보와는 무관하게 처리되는 경향을 가지고 있다. 또한 기준영상에 존재하는 특징정보가 다른 영상에서 존재하지 않거나 추출되지 않는 경우에는 상호간에 잘못된 정합(false-match)이 존재하며 이를 해결하기 위하여 부가적인 제한조건(constraints)를 두어 그 근처에서 해를 찾아야 한다. 제한조건은 일반적으로 물체의 표면에서 정의될 수 있는 물리적인 성질에서 타당성 있는 가정하에서 유도된다. 대표적인 제한조건으로는 유일성(uniqueness), 연속성(continuity), 순서성(ordering) 조건 등이 있다. 특징기반 방법은 에지(edge), 윤곽(contour, boundary) 등과 같은 특징정보를 정합하기 때문에 변이의 정확도가 높지만 조밀한 변이도는 구할 수 없는 단점을 가지고 있다. 이에 대하여 영역기반 방법은 국부적으로 작은 영상영역을 정합시킴으로써 조밀한 변이도를 구할 수 있지만 정합시키는 영상영역내의 모든 지점의 변이가 서로 유사하다는 가정을 하여야 하므로 경계부분에서 많은 오차가 나타난다. 영역기반 방식의 일반적인 처리방법은 최대 상관도(cross-correlation)나 최소 SSD(sum of squared difference)을 이용하는 것이다. 그러나 안정된 정합성능을 가지기 위해서는 영역의 크기를 최적화할 필요가 있다. 또한 상관관계 계를 계산하고 이들 중에서 최대화되는 영역을 찾는 정합과정은 설정된 전영역에서 처리되어야 하므로 많은 계산량이 요구되어 전체적인 처리속도가 떨어진다. 이러한 문제를 해결하는 방안으로 계층구조적 방법(hierarchical approach)을 택하지만 저해상도(coarse level)의 변이정보가 하위로 전달되기 때문에 고해상도(finer level)에서 다른 해를 찾거나 찾지 못하는 경우가 발생한다.

이상에서 논한 바와 같이 두가지의 접근방식에는 서로 장단점을 가지고 있으며, 상호간의 방식을 결합하는 방법이 제시되고 있으나 처리량에 비해 성능향상은 미흡하다고 보여진다. 스테레오 비전 시스템의 처리

결과는 영상의 전영역에 대한 변이도가 요구되고 있으므로 영역기반에 대한 연구가 활성화되고 있다고 보여진다. 이러한 과정에서 나타나는 중요한 관점은, "1. 정합척도를 어떻게 계산할 것인가? 2. 정합시킬 영역의 크기(support region)를 얼마로 할 것인가? 3. 후보 영역 중에서 최적의 정합점을 어떻게 찾을 것인가?" 하는 것이다. 정합척도는 두지점의 유사성을 결정하는 것으로 국부적인 화소단위(pixel unit)나 조그만 영역 상에서 결정된다. 정합 방법으로는 absolute intensity difference^[2,3], squared intensity difference^[4,5], binary pixel^[6], edges^[7], filtered image^[8-10], gradient^[11,12] 등이 있다. 정합척도는 주로 영상의 밝기차이, 밝기차이의 절대치, 밝기차이의 제곱, 상관도 등을 사용하며, 최적의 정합점은 이러한 정합척도를 최소화(minimizing) 혹은 최대화(maximizing) 과정을 거쳐 최상의 정합(best matching)을 찾는다. 이러한 과정에서 정합척도는 각 지점에 대하여 처리하는 것보다 그 주위의 정보도 함께 이용하는 것이 정합정도를 높일 수 있다. 따라서 주위의 정보를 모으는 방법이 강구되어야 하는데 보통 사각창(square window)을 주로 이용하며, Gaussian convolution^[12], adaptive window^[3], disparity gradient^[13], Prazdny's coherence principle^[14], diffusion process^[5,15] 등을 사용한다. 상관도나 SSD에 근거한 스테레오 정합방법에서는 응용분야에 따라 고정된 영역크기를 사용한다. Barnard와 Fischer^[16]가 지적하였듯이 상관도 혹은 SSD를 이용한 스테레오 정합에서 중요한 과제는 정합에 필요한 충분한 밝기변화를 포함할 만큼 크고 투시변형(projective distortion)의 영향을 받지 않을 만큼 작은 창크기를 결정하는 것이라 했다. 만약 영역의 크기가 너무 작아 충분한 밝기변화를 포함하지 못하면 잡음의 영향을 받아 변이추정(disparity estimate)이 부실해지며, 반대로 너무 크게 설정하면 변이의 변화에 무디어져 폐색영역(occlusion region) 등과 같은 세밀히 관찰해야 할 영역에서 무디어지는(blurring) 결과를 낳게 된다.

이러한 이유로 하여 영역기반 방식에서는 국부적인 밝기변화와 변이변화를 고려하여 최적의 정합척도를 모으는 영역의 크기를 결정할 수 있어야 한다. 최적의 정합척도 영역의 크기를 결정하기 위해서 "1. 효율성: 컨볼루션과 같은 간단한 동작으로 모을 수 있어야 한다.^[3] 2. 병렬성: 병렬처리 형태의 구조를 가지도록

하여야 한다.^[15] 3. 적응성: 창 의 크기를 조절할 수 있어야 한다.^[3,15] ” 와 같은 특성을 가지도록 그 처리방법을 강구하여야 한다.

본 연구에서는 이러한 효율성, 병렬성 및 적응성을 가지고 있는 확산망^[17-20]을 이용하여 스테레오 정합의 영역크기 조절문제를 해결하고자 하였다. 확산망은 4개의 이웃 점과의 연결세기로 구성되며, 연결세기를 통한 처리과정을 반복적으로 수행함으로써 정합시킬 영상영역의 크기를 쉽게 조절할 수 있다. II장에서는 제시된 확산망을 가지고 그 특징들에 대해 살펴보고, III장에서는 확산망을 가지고 구성된 실제 스테레오 비전시스템에 관하여 논하며, IV장에서는 스테레오 비전시스템에 합성영상 및 자연영상을 적용하였을 때의 실험결과를 보이고, 이를 통해 확산망을 이용하여 정합 척도를 모으는 스테레오 비전시스템의 성능을 살펴볼 것이다.

II. 확산망

동물 신경망의 수용영역에서 신경세포에 외부자극이 인가되면 그 자극은 주변의 신경세포들에게 전달되면서 주변세포로 퍼지게 된다. 한 신경세포에 입력자극이 인가되었을 때 인접한 다른 신경세포들에게로 자극의 일부가 전달되고 자신도 자극의 영향을 받는 형태로 계속 확산되면 자극은 주변으로 점점 더 퍼지고 그 결과로 반응의 분포는 가우스 분포에 근접하게 된다. 이때 확산에 의한 분포는 확산이 거듭됨에 따라 분산이 더 큰 가우스 함수의 모양이 된다. 본 논문에서는 먼저 이러한 특징을 가진 확산망을 확산을 진행하는 방법에 따라 확산의 방향성을 가지지 않고 일률적으로 확산이 발생하는 등방성 확산(isotropic diffusion)과 확산을 할 때 영상의 특징을 고려한 방향성을 가진 비등방성 확산(anisotropic diffusion)으로 나누어 제안하고, 각각의 특징을 논의한다.

1. 등방성 확산망

등방성 확산망은 확산의 방향이 특정 방향으로 치우치지 않고 일률적으로 모든 방향으로 일어나는 것으로 다음과 같은 식으로 정의될 수 있다.

$$\alpha \nabla^2 E = \frac{\partial E}{\partial t} \tag{1}$$

여기서 α 는 확산계수이고, E 는 입력을 나타낸다. 이

것을 이산적인 형태로 변형하면

$$E(i, j) \leftarrow (1-4\alpha)E(i, j) + \alpha \sum_{(k, l) \in N_4} E(i+k, j+l)$$

와 같이 표현된다. 여기서 N_4 는 $\{(-1,0),(1,0),(0,-1),(0,1)\}$ 로서 4방향의 이웃점을 나타내며 $(1-4\alpha) > 0$ 관계를 가져야 한다. 이러한 관계식을 반복 알고리즘으로 변환하면

$$E^{n+1}(i, j) \leftarrow (1-4\alpha)E^n(i, j) + \alpha \sum_{(k, l) \in N_4} E^n(i+k, j+l) \tag{2}$$

로 표현된다. 여기서 n 은 확산횟수이다. 즉, 2차원 확산과정을 반복적으로 수행할 때마다 자기 자신의 정보로부터 $(1-4\alpha)$ 의 비율과 인접 4개의 셀(cell)로부터 α 만큼의 정보가 모여져서 정보가 새롭게 만들어 진다. 그림 1은 $\alpha=0.25$ 일때의 입력이 임펄스인 경우에 1차원에서의 확산을 보여주는 것으로 반복횟수를 거듭할수록 그 분포가 가우스분포에 가까워짐을 알 수 있다. 이때 반복횟수에 대한 분산 정도는 $\sigma^2 = 2an$ 과 같은 관계식을 가진다. 이것은 확산횟수와 확산계수를 조절함으로써 창 의 크기 즉, 정보를 모으는 셀의 범위를 간단하게 조절할 수 있음을 보여주고 있다. 그러나 초기영상에 관계없이 확산횟수를 거듭할수록 확산이 선형적으로 계속되는 선형확산법(linear diffusion)은 초기영상의 불연속성이 무시되면서 확산을 하는 경향이 있다. 그래서 초기 영상을 계속적으로 보여주면서 확산을 하는 방법으로 다음과 같은 확산방정식을 고려할 수 있다.

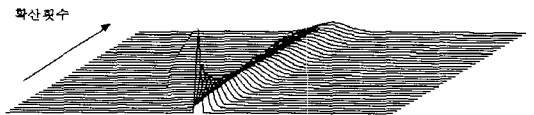


그림 1. 단위 임펄스 함수에서의 확산
Fig. 1. Diffusion of unit impulse.

$$\alpha \nabla^2 E + \beta E^0 = \frac{\partial E}{\partial t} \tag{3}$$

여기서 E^0 는 초기입력영상이고, β 는 초기입력영상을 고려하기 위한 가중치이다. 이것을 이산적인 수식으로 표현하면

$$E(i, j) \leftarrow (1-4\alpha)E(i, j) + \alpha \sum_{(k, l) \in N_4} E(i+k, j+l) + \beta E^0$$

와 같이 표현될 수 있다. 이것을 다시 반복 알고리즘

으로 표현하면 다음과 같이 된다.

$$E^{n+1}(i, j) \leftarrow (1-4\alpha)E^n(i, j) + \alpha \sum_{(k, \eta) \in N_i} E^n(i+k, j+\eta) + \beta E^0 \quad (4)$$

식(4)와 같이 초기영상을 고려하기 위한 β 가중치를 넣음으로써 확산이 초기영상에 무관하게 선형적으로 확산이 일어나는 선형 확산망과 달리 초기영상에 크게 벗어남이 없이 확산시킬 수 있는 비선형확산 (nonlinear diffusion)망을 구현할 수 있다. 이런 선형 확산망과 비선형 확산망은 그림 2와 같이 간단하게 구현될 수 있으며, 이때 두 확산망은 인접한 4개의 셀로의 확산 정도가 일률적으로 같은 등방성 확산으로 동일한 확산계수를 가지고 있다 선형 확산망은 초기영상을 고려하지 않으므로 β 값이 0으로 되고, 비선형 확산망은 초기영상을 고려하므로 β 의 값을 가진다.

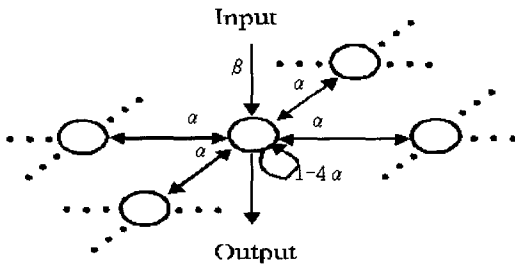


그림 2. 2차원에서의 확산망 구조
Fig. 2. 2-D diffusion network.

2. 비등방성 확산망

비등방성 확산 방정식은

$$\nabla(\alpha(x, y, t)\nabla E) = \frac{\partial E}{\partial t} \quad (5)$$

와 같이 표현된다. 여기서 $\alpha(x, y, t)$ 는 확산함수이다. 이 확산함수는 등방성 확산망의 확산계수와 달리 영상에서 화소마다 값이 달라진다. 특히 두 영역이 부딪치는 경계부분을 알 수 있다면 영역내에 있는 화소의 확산함수는 1에 가깝게 하고, 경계부분을 지나는 다른 영역과의 확산함수는 0에 가깝게 함으로써 확산이 진행됨에 따라 동일한 영역에서만 정보를 모으게 된다. 따라서 확산함수를 경계값이 추정가능한 인자로 하는 함수로 선택한다면 확산이 경계부분을 넘지않고 동일한 영역에만 확산을 주로 발생시킬 수 있다. 본 논문에서는 경계값을 추정하는 방법으로 아래와 같이 영상의 밝기의 기울기 값을 사용하였다.

$$\alpha(x, y, t) = f(\|\nabla E(x, y, t)\|)$$

위에서 함수 f 는 경계부분에서 확산함수를 0에 가깝게 하기 위해 경계부분에서 큰 값을 가지는 밝기의 기울기 값 $\|\nabla E(x, y, t)\|$ 에 단조감소적인 함수를 적당히 선택하면 된다. 본 논문에서는 다음과 같은 단조감소 함수를 정의하여 확산함수를 구하였다.

$$\alpha(x, y, t) = f(\|\nabla E(x, y, t)\|) = \frac{1}{1 + (\frac{\|\nabla E(x, y, t)\|}{K})^2}$$

이때 K 는 상수이다. $\alpha(x, y, t)$ 가 일정한 값을 가지면 등방성 확산망이 된다. 위의 확산함수를 가지고 4방향 이웃 점에 대하여 이산적인 반복 형태로 표현하면

$$E^{n+1}(i, j) = (1 - \sum_{(k, \eta) \in N_i} \alpha^n(k, \eta))E^n(i, j) + \sum_{(k, \eta) \in N_i} \alpha^n(k, \eta)E^n(i+k, j+\eta) \quad (6)$$

과 같이 표현된다. 식 (6)식에 사용된 확산함수를 영상의 초기치에 의존하도록 하면 다음과 같이 확산함수를 영상의 각 화소마다 초기에 구하여 확산을 진행할 때 초기영상을 고려한 확산함수를 사용하게 된다.

$$\alpha(x, y) = f(\|\nabla g(x, y)\|)$$

여기서 $g(x, y)$ 는 초기영상에 의존하는 함수로써 $g(x, y) = E(x, y, 0)$ 등으로 둔다. 초기영상을 통해 만들어진 확산함수를 가지고 확산을 진행하는 비등방성 확산망인 가중 확산망은 식 (6)과 마찬가지로 다음과 같은 이산적인 반복 형태로 표현되어진다.

$$E^{n+1}(i, j) = (1 - \sum_{(k, \eta) \in N_i} \alpha(k, \eta))E^n(i, j) + \sum_{(k, \eta) \in N_i} \alpha(k, \eta)E^n(i+k, j+\eta) \quad (7)$$

이상과 같이 확산망은 한 셀의 이웃 4방향에서 정보를 모으는 방법에 따라 크게 두가지로 분류하면, 확산이 일어날 때 일률적으로 모으는 식 (2)와 (4)의 등방성 확산망과 영상의 국부적인 특징을 가지고 동일한 영역에 존재하는 정보만을 모으려는 식 (6)과 (7)의 비등방성 확산망으로 제안할 수 있다. 그림 3은 밝기 정보가 명확한 합성영상에 10%의 가우시안 잡음을 첨가하였을 때의 확산의 실험을 보여주고 있다. 먼저 그림 3(a)는 합성영상이며 그림 3(b)는 잡음이 첨가된 합성영상을 보여주며 그림 3(c),(d),(e),(f)는 선형 확산법, 비선형 확산, 비등방성 확산법 그리고 초기영상을 고려한 비등방 확산인 가중확산법에서의 각각의 확

산결과를 보여주고 있다. 이때 확산 반복횟수는 150으로 하였다. 위에서 언급하였듯이 그림 3(c)의 선형 확산법은 이웃한 4방향에서 초기 영상에 관계없이 일률적인 비율로 정보를 모으므로 영역이 변하는 경계부분에서는 정보가 무디어져 영상의 특징적인 정보를 손상시키는 확산을 하는 경향을 보이고 있다. 그림 3(d)의 비선형 확산법은 초기영상 및 영상의 특징을 어느 정도 고려하면서 확산이 진행함을 보여주고 있다. 그림 3(e)와 (f)의 비등방성 확산법과 가중확산법은 확산이 진행될 때 영상의 국부적인 정보를 고려한 확산함수를 통해 주위정보를 모으므로 영역이 달라지는 경계점을 넘어서는 부분의 정보는 모이지 않고 같은 영역 안에

서만 주위 정보를 모으므로 경계부분이 강화되면서 확산이 진행된 것을 볼 수 있다. 전체적으로 보면 식 (6),(7)의 비등방성확산법과 식 (4)의 비선형 확산법은 경계를 강화하면서 주위 정보를 모으는 능력이 뛰어난 것을 보여준다. 표 1은 그림 3(a)의 원영상과 10%의 잡음을 첨가한 후 확산을 통해 잡음이 제거된 영상과의 RMS 에러를 보여주고 있다. 비등방 확산법과 가중확산법이 경계값을 강화하면서 확산을 통해 주위 정보를 모아서 잡음을 제거하여 등방성 확산법에 비해 에러가 적음을 알 수 있다.

표 1. 확산망의 특성비교

Table 1. Performance of Diffusion Network.

Method	RMS Error [intensity]
Linear Diffusion(eq. (2))	0.110
Nonlinear Diffusion(eq. (4))	0.077
Anisotropic Diffusion(eq. (6))	0.061
Weighted Diffusion(eq. (7))	0.065

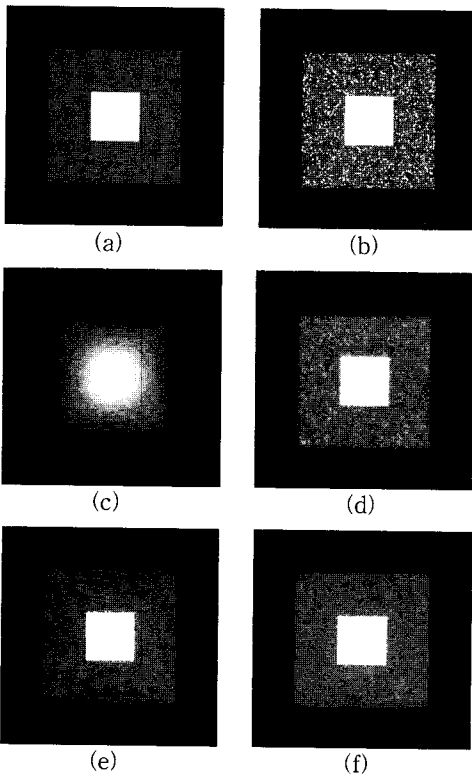


그림 3. 확산방법에 따라 확산된 영상, 확산횟수 150번

(a) 원영상 (b) 잡음이 첨가된 영상 (c) 선형 확산(식 (2)) (d) 비선형 확산(식 (4)) (e) 비등방성 확산(식 (6)) (f) 가중 확산(식 (7))

Fig. 3. Diffused images by each diffusion algorithm.

(a) original image (b) original image with 10% Gaussian noise (c) linear diffusion(eq. (2)). (d) nonlinear diffusion(eq. (3)) (e) anisotropic diffusion(eq. (6)).(e) weighted anisotropic diffusion(eq. (7)).

III. 확산망을 이용한 스테레오 비전 시스템

스테레오 정합의 전체적인 구성도는 크게 세부분으로 나눌 수 있다. 먼저 좌,우 영상에서 정합 척도를 구하는 부분, 정합척도를 모으는(aggregating)부분, 그리고 모아진 정합척도를 가지고 변이정보를 구하는 부분으로 나누어 진다.

정합척도는 일반적으로 두 개의 영상에서 다음과 같이 표현된다.

$$E(x, y, 0; d) = f(I_L(x + d, y) - I_R(x, y)) \tag{9}$$

여기서, I_L 과 I_R 은 각각 좌우영상의 밝기정보이며, d 는 변이, f 는 두영상 밝기에 대한 유사성을 나타내는 함수로서 본 연구에서는 제곱차(squared difference)를 사용하였다. 또한 $E(x, y, 0; d)$ 는 정합척도를 모으기 이전의 초기상태이다. 정합척도를 모으는 대표적인 방법으로서 컨볼루션으로 정합척도를 모으는 범위는 창크기로 결정된다. 일반적으로 고정된 크기를 사용한다. 정합척도를 모은 결과를 $E(x, y, t; d)$ 라 할 때 (x, y) 지점의 변이는

$$d(x, y) = \min_{d \in D} \{E(x, y, t; d)\} \tag{10}$$

로 표현된다. 여기서 D 는 변이범위로 $D = [dmin, dmax]$ 로 제한된다. 즉, (x, y) 지점의 변이는 변이 범위내에서 모아진 정합척도중에서 최소값을 가지는 위치의 변이 d 가 된다는 것이다.

고정된 창을 가지는 SSD 알고리즘은 확실히 정합척도를 모으는 효과를 나타낸다. 하지만 앞에서 언급하였듯이 잡음에 의한 변이의 변형과 깊이(depth)의 흐릿해짐 사이에는 상호관계(trade-off)가 존재하게 된다. 이를 해결하기 위한 방안으로 Gaussian과 같은 non-uniform window^[12]를 사용하여 창의 가중치를 달리해서 모으는 정도를 조정하거나 adaptive window^{[2][3]}를 사용하여 창의 변이를 추정하여 유사한 변이만을 가지도록 창의 크기를 조절하기도 한다. 그러나 창의 가중치를 다르게 하여 정합의 성능은 개선할 수가 있으나 근본적인 상관관계는 해결되지 않으며, 창의 크기를 조절할 수밖에 없다. 또한 영상의 성질에 따라 창의 크기를 다르게 적용할 때 상관관계 문제는 어느 정도 해결할 수 있으나 하드웨어적으로 실현하기가 곤란하다. 따라서 창의 가중치를 달리 하면서 하드웨어적으로 간단히 구성될 수 있고, 또한 창의 크기를 가변시키면서 정합척도를 모으는 방법이 필요하게 된다. 본 논문에서는 스테레오 정합의 성능에 크게 영향을 미치는 위의 문제를 해결하기 위해 확산망을 이용하였다. 그림 4는 확산망을 이용하여 창의 크기를 조절할 수 있는 스테레오 비전시스템 모델을 나타낸 것이다. II장에서 언급하였듯이 확산망은 창의 가중치를 가질 수 있으며 또한 창의 크기도 반복확산 횟수를 조절함으로써 변화시킬 수 있다. 확산망은 정보를 모으는 방법에 따라 크게 등방성 확산망과 비등방성 확산망으로 나누었다. 스테레오 정합에서 등방성 확산망은 초기 정합척도에서 확산을 통하여 정합척도를 모을 때 방향성을 고려하지 않고, 이웃 4방향에서 일률적으로 똑같은 비율로 정합척도를 모으는 성질이 있고, 비등방성 확산망에서는 초기 정합척도에서 확산을 통하여 정합척도를 모을 때 정합척도에서 인접한 경계점을 넘어서 변이가 달라지는 부분에서는 확산합수가 0에 가까우므로 변이가 달라지는 영역에서는 정합척도를 모으지 않게 된다. 따라서 비등방성 확산망에서는 정합척도에서 같은 변이를 갖는 영역에서 정합척도를 주로 모으고 다른 변이에서는 정합척도의 정보를 모으지 않는 성질을 가진다. 또한, 정합척도를 모으는 창의 모양이 영상의 화소마다 달라져 Kanade가

제안한 adaptive window^[3]의 효과를 쉽게 구현할 수 있다. 그림 5는 변이 정보가 물체에 해당하는 부분 df 와 배경에 해당하는 db 로 구분될 때 위치 P_0, P_1, P_2, P_3 에서 등방성 확산망을 이용하는 경우와 비등방성 확산망을 이용하여 정합척도를 모으는 경우에 창의 모양의 변화를 보인 것이다. 동일한 변이정보를 갖는 영역의 중심에 위치한 P_0 와 같은 지점에서는 등방성, 비등방성 확산망 모두 모든 방향으로 확산을 발생시켜 방향성을 가지지 않고 주위의 정합 척도를

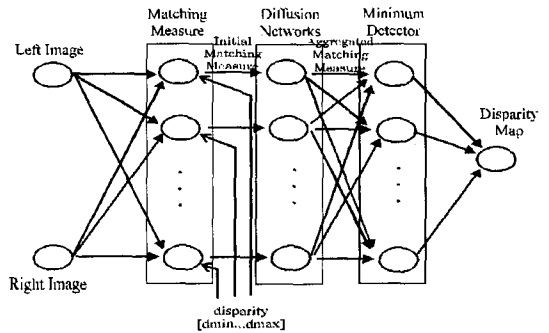


그림 4. 확산망을 이용한 스테레오 비전 시스템 모델
Fig. 4. Stereo vision system using diffusion networks.

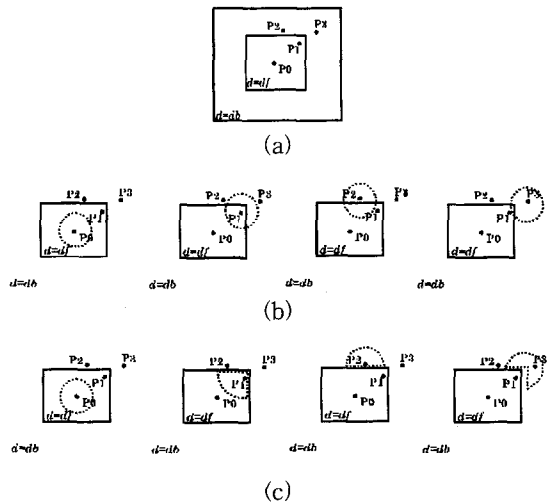


그림 5. 확산을 할 때 창의 모양 변화
(a) 변이 정보 (b) 등방성 확산망에서의 창의 변화 (c) 비등방성 확산망에서의 창의 변화
Fig. 5. Window shape by diffusion.

(a) Disparity information. (b) Window shape by isotropic diffusion (c) Window shape by anisotropic or wighted diffusion.

모이지만, P_1, P_2, P_3 지점에서는 변이가 달라지는 불연속 지점에 위치하므로 비등방성 확산망에서는 확산이 같은 변이를 가진 영역 쪽으로 발생하여 정합척도를 모으는 창 모양이 위치에 따라 변하게 된다.

IV. 실험 및 결과

본 논문에서 선형 확산망, 비선형 확산망, 비등방성 확산망 그리고 가중확산망을 이용하여 초기 정합척도로부터 정합척도를 모아 변이를 추출하는 실험을 하였다. 실험에서는 먼저 256×256 크기의 변이의 불연속 정도가 뚜렷한 합성 영상을 이용해 실험하였고, 그리고 실제 자연영상에도 적용하여 실험하였다. 스테레오 영상에서 변이를 추출하기 위한 전체 알고리즘은 III장에서 나타낸 바와 같이 다음과 같은 알고리즘을 사용하였다.

1. 각 변이에서 스테레오의 좌,우 두영상을 가지고 초기 정합척도를 다음과 같이 구한다.

$$E(x, y, 0; d) = (I_L(x+d, y) - I_R(x, y))^2, \quad d = d_{\min}, \dots, d_{\max}$$

2. 각각의 변이에 따른 초기 정합척도를 확산망의 입력으로 하여 식 (2), 식 (4), 식 (6) 및 식 (7)에서 정의한 확산망을 통해 정합척도를 모은다.
3. 영상의 각 지점에 대하여 최소의 확산정합척도를 나타내는 변이를 추정한다.

먼저 실험에 사용된 합성 영상은 그림 6의 rds(random dot stereogram)영상과 그림 8의 ramp 스테레오 영상을 가지고 실험을 하였다. rds 합성 영상은 변이가 0, 5, 10으로 변이정보가 뚜렷한 스테레오 영상이다. 그림 7은 각 확산망을 가지고 실험하였을 때의 변이도를 보여주고 있다. 최종 검출된 변이도는 상대적인 거리를 밝기로 표현한 것으로 가장 밝은 부분이 변이가 가장 큰 10으로 카메라와 물체와의 거리가 가장 가까운 부분이며, 가장 어두운 부분이 변이가 0인 카메라와 물체와의 거리가 가장 먼 부분이다. 확산 횟수는 40으로 고정하였으며, 비등방확산에서 확산함수를 구하기 위한 단조감소 함수의 최적의 상수 K값을 설정하기 위해 본 논문에서는 III장에서 언급된 가우시안 잡음이 첨가된 합성영상의 확산된 영상에서 RMS 에러를 최소화 시키는 K값으로 10을 설정하였다.

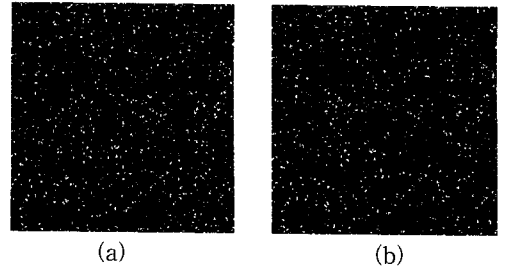
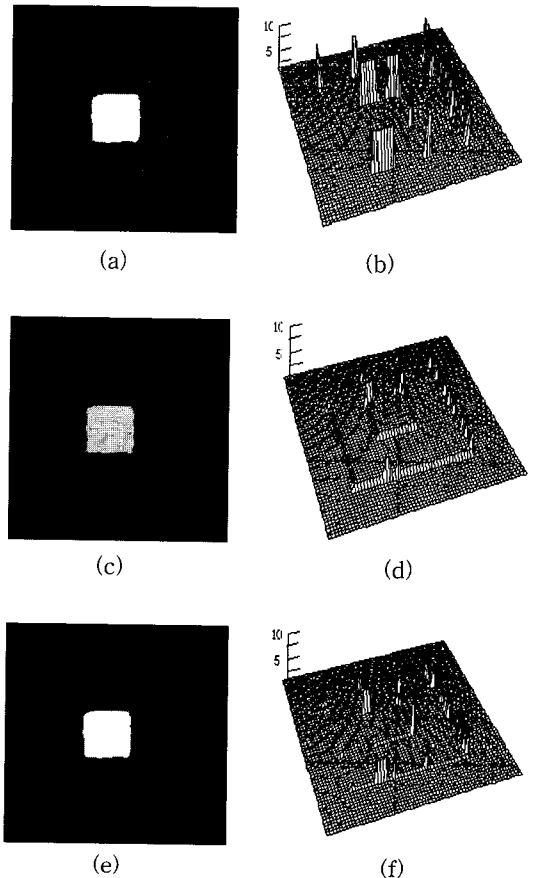


그림 6. 합성 영상 rds
(a) 좌 영상 (b) 우 영상
Fig. 6. Synthetic stereo image.
(a) Left image. (b) Right image.

표 2. 확산망에 따른 합성 영상의 RMS 에러
Table 2. RMS Error by diffusion methods.

Method	RMS Error [pixel]
Linear Diffusion(eq. (2))	0.144
Nonlinear Diffusion(eq. (4))	0.123
Anisotropic Diffusion(eq. (6))	0.088
Weighted Diffusion(eq. (7))	0.089



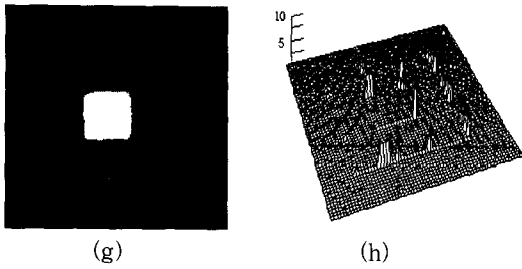


그림 7. 합성 영상의 스테레오 정합에 따른 변이도와 실제 변이도와의 에러맵
 (a) 선형 확산 (식 (2)) (b) 선형 확산에 의한 에러맵 (c) 비선형 확산 (식 (4)) (d) 비선형 확산에 의한 에러맵 (e) 비등방성 확산 (식 (6)) (f) 비등방성 확산에 의한 에러맵 (g) 가중확산 (식 (7)) (h) 가중확산에 의한 에러맵

Fig. 7. Depth map by stereo matching of synthetic image and Error map.

(a) Linear diffusion (eq. (2)) (b) Error map by linear diffusion (c) Nonlinear diffusion (eq. (4)) (d) Error map by nonlinear diffusion (e) Anisotropic diffusion (eq. (6)). (f) Error map by anisotropic diffusion (g) Weighted diffusion (eq. (7)) (h) Error map by weighted diffusion.

그림 7은 III장에서 언급되었던 확산망에 따른 변이도와 실제 변이도와의 에러맵을 나타낸 것이다. 그림 7(a)와 (b)는 선형 확산법에 의한 변이도와 에러맵으로써 변이가 동일한 영역 안에서는 에러가 거의 없지만 변이가 달라지는 불연속 부분에서는 에러가 많이 발생함을 볼 수 있다. 이는 선형 확산법이 확산을 통해서 초기 정합척도로부터 정합척도를 모을 때 변이가 달라지는 불연속 부분 주위에서는 변이가 다른 영역의 정합척도까지 모았기 때문에 발생한 것이다. 그림 7(c)와 (d)의 비선형 확산법은 선형 확산법에 비해서 초기 정합척도 상태를 계속 보여주기 위해 그것에 크게 벗어나지 않으려는 가중치를 두었으므로 불연속 부분에서도 에러가 비교적 적음을 볼 수 있다. 그림 7(e)와 (g)는 비등방성 확산과 가중확산으로 영상의 화소마다 영상의 특징을 반영한 확산함수를 가지고 확산하였으므로 불연속 부분에서도 위의 선형, 비선형 확산에 비해 에러가 적음을 볼 수 있다. 표 2는 정합결과로 나타난 변이도와 실제 변이도 사이의 RMS 에러를 나타낸 것이다.

그림 8은 실험에 사용된 다른 합성영상인 ramp 합성영상과 실제 변이도를 나타낸 것이다. 이 합성영상은 변이가 2, 6의 뚜렷한 변이도를 가지고 있다. 그림 9는 이 합성영상을 가지고 등방확산의 비선형 확산과 비등방 확산의 가중확산을 가지고 실험하였을 때의 확산횟수에 따른 정합척도의 변화를 보인다.

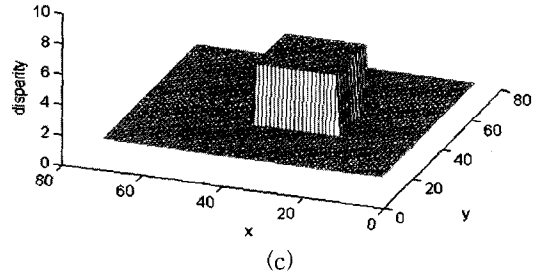
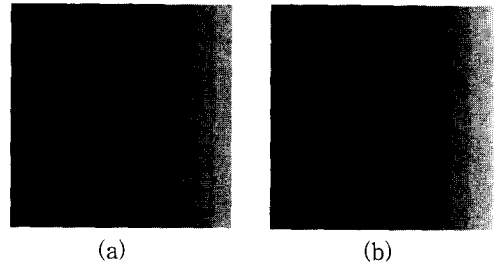


그림 8. ramp 스테레오 영상과 실제 변이도
 (a) 좌 영상 (b) 우 영상 (c) 실제 변이도
 Fig. 8. Ramp stereo image and true disparity.
 (a) Left image (b) Right image (c) True disparity

그림 10은 자연영상인 펜타곤(Pentagon)영상을 각 확산망에 적용하여 정합 실험을 했을 때 구해진 변이도와 고정된 창 크기를 갖는 SSD알고리즘으로 정합척도를 모아서 얻은 변이도를 보이고 있다. 자연영상에서도 합성영상을 가지고 실험하였을 때와 같은 특징을 보이고 있으며, 특히 비선형 확산은 식 (3)에서와 같이 초기 정합상태를 고려한 β 가중치를 이용하고 가중 확산은 식 (7)에서와 같이 초기 정합상태에 의존하는 확산함수를 이용하여 정합척도를 모으므로써 세밀한 변이도를 얻을 수 있었다. 이상에서 서술한 바와 같이 제안된 4가지 확산을 통해 정합척도를 모으는 방법이 모두 SSD알고리즘보다 우수한 결과를 얻었으며 비선형확산, 비등방성확산, 그리고 가중확산에서의 결과가 보다 우수함을 볼 수 있다.

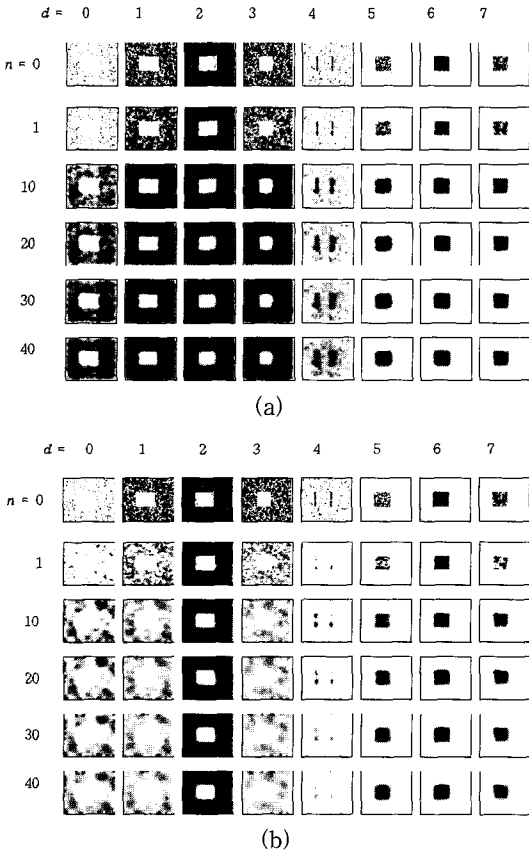


그림 9. 확산횟수에 따른 정합척도의 변화
 (a) 비선형 확산인 경우 (b) 가중확산인 경우
 Fig. 9. Change of matching measure by diffusion iteration.
 (a) Nonlinear diffusion (b) Weighted diffusion

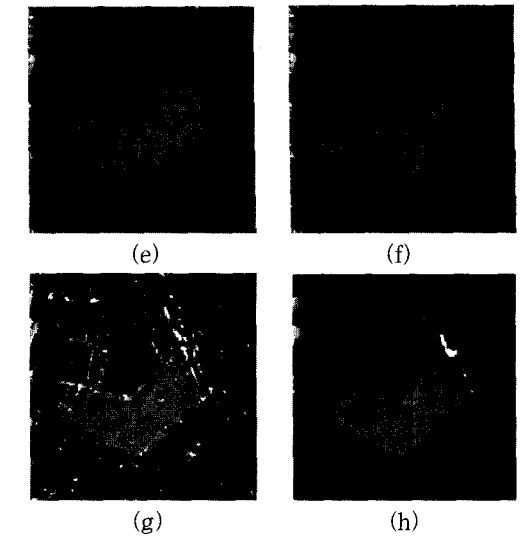
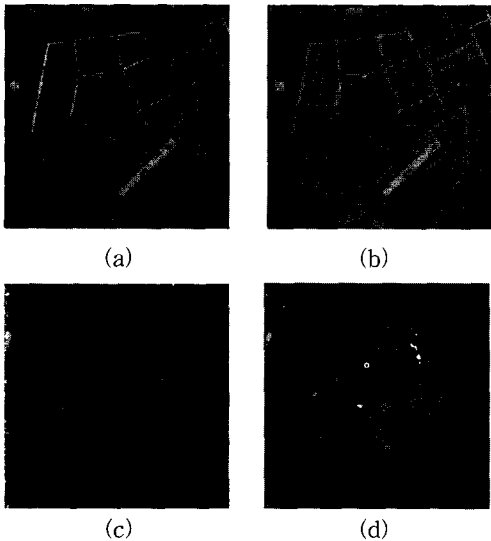


그림 10. 펜타곤 스테레오 영상의 변이 추정 결과
 (a) 좌 영상 (b) 우 영상 (c) 선형 확산
 (식 (2)) (d) 비선형 확산 (식 (4)) (e) 비
 등방성 확산 (식 (6)) (f) 가중확산 (식 (7))
 (g) 5×5 SSD (h) 13×13 SSD
 Fig. 10. The disparity map of Pentagon stereo
 image.
 (a) Left image (b) Right image (c) Linear
 diffusion (eq. (2)) (d) Nonlinear diffusion
 (eq. (4)) (e) Anisotropic diffusion (eq.
 (6)) (f) Weighted diffusion (eq. (7)) (g)
 5×5 SSD. (h) 13×13 SSD

V. 결론

본 논문에서는 영역기반 스테레오 정합에서 정합 성능을 크게 좌우하는 중요한 문제로서 비교해야 할 영상영역의 크기를 최적화하는 방안으로 확산망을 제안하였다. 제안된 확산망은 크게 선형, 비선형 확산망과 같은 등방성 확산망, 비등방성 확산망, 그리고 초기영상이 고려된 가중확산망이며, 이러한 확산망은 스테레오 정합과정에서 주위의 정보를 모으는 데 적용하였다. 확산망에 의한 스테레오 정합결과는 기존의 널리 사용되는 알고리즘인 SSD를 적용하였을 때보다 우수함을 실험결과를 통하여 알 수 있었다. 확산망을 사용함에 따라 효율적으로 창 크기를 조절할 수 있었을 뿐만 아니라, 확산동작이 이웃된 셀과의 연결만으로 이루어지므로 차후 하드웨어로 실현할 때 보다 효율적이라 하겠다. 본 논문에서는 확산횟수를 고정시켰으나 확산 횟수를 자동적으로 결정하는 알고리즘에 대한 연구가 계속 진행되어야 하겠다.

감사의 글

※ 본 논문은 부산대학교 컴퓨터 및 정보통신연구소의 지원에 의해 수행된 바 이에 감사를 드립니다.

참 고 문 헌

- [1] U. R. Dhond and J. K. Aggarwal, "Structure from stereo - a review," *IEEE Trans. on systems, Man, and Cybernetics*, vol. 19, no. 6, pp. 1189-1510, Nov./Dec. 1989.
- [2] M. Okutomi and T. Kanade, "A locally adaptive window for signal matching," *Int. J. of Computer Vision*, vol. 7, no. 2, pp. 113-162, Apr. 1992.
- [3] T. Kanade and M. Okutomi, "A stereo matching algorithm with an adaptive window: theory and experiment," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 16, no. 9, pp. 920-932, Sep. 1994.
- [4] L. H. Matthies. R. Szeliski and T. Kanade, "Kalman filter-based algorithms for estimating depth from image sequences," *Int. J. of Computer Vision*, vol. 3, pp. 209-236, 1989.
- [5] D. Scharstein and R. Szeliski, "Stereo matching with non-linear diffusion," *Proc. Computer Vision and Pattern Recognition*, pp. 343-350, California. Jun. 1996.
- [6] D. Marr and T. Poggio, "Cooperative computation of stereo disparity," *Science*, vol. 94, pp. 283-287, Oct. 1976.
- [7] H. H. Baker, "Edge based stereo correlation," *Image Understanding Workshop*, pp. 168-175, April 1980.
- [8] D. Marr and T. Poggio, "A computational theory of human stereo vision," *Proc. of the Royal Society London*, vol. B-204, pp. 301-328, 1979.
- [9] D. Jepson and K. Tsotsos, "Techniques for disparity measurement," *CVGIP: Image Understanding*, vol. 53, no. 1, pp. 14-30, Jan. 1991.
- [10] D. G. Jones and J. Malik, "A computational framework for determining stereo correspondence from a set of linear spatial filter," *2nd European Conf. on Computer Vision*, pp. 397-410, Italy, May 1992.
- [11] P. Seitz, "Using local orientation information as image primitive for robust object recognition," *SPIE Visual Com. and Image Processing IV*, vol. 1199, pp. 1630-1639, 1989.
- [12] D. Scharstein, "Matching images by comparing their gradientfields," *12th Int. Conf. on Pattern Recognition*, vol. 1, pp. 572-575, Israel, Oct. 1994.
- [13] S. B. Pollard. J. E. W. Mayhew. and J. P. Frisby, "PMF : A stereo correspondence algorithm using a disparity gradient limit," *Perception* vol. 14. pp. 449-470. 1985.
- [14] K. Prazdny, "Detection of binocular disparities," *Biological Cybernetics*, vol. 52, no. 2 pp. 93-99, 1985.
- [15] R. Szeliski and G. Hinton, "Solving random-dot stereograms using the heat equation," *IEEE Conf. Computer Vision and Pattern Recognition*, pp. 284-288, San Francisco, Jun. 1985.
- [16] S. T. Barnard and M. A. Fischer, "Stereo vision," in *Encycloperdia of Artificial Intelligence*, pp. 1083-1090, New York : Jon Wiley, 1987.
- [17] 권률, 남기곤, 윤태훈, 김재창, "확산뉴런망을 이용한 밝기변화 추출," *전자공학회논문지*, 제 29 권 B편, 제11권, pp. 85-92, 1992
- [18] 권률, 남기곤, 윤태훈, 김재창, "확산뉴런망을 이용한 영상처리," *전자공학회논문지*, 제30권 B편, 제1권, pp. 90-99, 1993
- [19] P. Perona and J. Malik, "Scale - Space and Edge Detection Using Anisotropic Diffusion," *IEEE Tans. Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629-640, July 1990.
- [20] P. Saint-Marc, J. Chen and G. Medioni, "Adaptive Smoothing : A General Tool for Early Vision," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol 13, no. 6, pp. 514-529, Jun. 1991.

저 자 소 개



李相璨(正會員)

1959년 9월 24일생. 1985년 2월 부산대학교 전자공학 공학사. 1985년 1월 ~ 1991년 6월 효성 T&C 컴퓨터사업부 근무. 1994년 2월 부산대학교 대학원 전자공학 공학석사. 1996년 2월 부산대학교 대학원

전자공학과 박사수료. 1997년 3월 ~ 현재 동의공업전문대학 영상정보처리과 전임강사. 주관심분야는 컴퓨터 비전임

南基坤(正會員) 현재 부산대학교 전자공학과 부교수

金在昌(正會員) 현재 부산대학교 전자공학과 교수

姜昌淳(正會員) 현재 동의공업전문대학 전자과 부교수

鄭斗泳(正會員)

1959년 11월 13일생. 1982년 2월 동아대학교 전자공학 공학사. 1986년 2월 부산대학교 대학원 전자공학 공학석사. 1986년 1월 ~ 1990년 8월 금성사 정보기기 연구소 근무. 1996년 2월 부산대학교 대학원 전자공학 박사수료. 1991년 3월 ~ 현재 거제전문대학 전자과 조교수. 주관심분야는 컴퓨터 비전임

李尙昱(正會員) 현재 경상대학교 정보통신공학과 교수