

시·공간 중복성을 고려한 다이내믹 프로그래밍 기반의 고속 변이 추정 기법

준회원 윤정환*, 배병규*, 박세환*, 정회원 송혁***,
종신회원 김동욱**, 정회원 유지상*

Fast Disparity Estimation Method Considering Temporal and Spatial Redundancy Based on a Dynamic Programming

Jung-hwan Yun*, Byung-kyu Bae*, Se-hwan Park* *Associate Members*,
Hyok Song*** *Regular Member*, Dong-wook Kim** *Lifelong Member*,
Ji-sang Yoo* *Regular Member*

요약

본 논문에서는 스테레오 정합에 사용되는 다이내믹 프로그래밍을 기반으로 영상의 시·공간적 중복성을 고려한 고속 변이 추정 기법을 제안한다. 영상 내 인접한 화소 간에 존재하는 높은 상관성을 이용하여 격행으로 다이내믹 프로그래밍 기반의 변이 추정을 수행하고 변이가 할당되지 않은 라인은 동일영역 내에서의 변이가 유사하다는 성질을 이용하여 주변의 변이벡터로 할당하였다. 또한 동영상의 변이 추정 시에는 프레임 간의 시간적 중복성을 고려한 알고리즘으로 수행시간을 크게 단축시킬 수 있었다. 모의 실험을 통하여 변이 정보를 이용한 중간시점 영상 생성 결과 블록매칭 알고리즘에 비하여 0.8~2.4dB의 PSNR(peak signal to noise ratio)이 증가하는 것을 확인하였다. 또한 일반적인 다이내믹 프로그래밍 기반의 변이 추정 방식과의 비교에서도 0.1dB가량의 PSNR이 증가하였으며 약 48~68%의 계산량이 감소하는 것을 확인 할 수 있었다.

Key Words : Disparity estimation, Dynamic programming, Temporal and spatial redundancies

ABSTRACT

In this paper, we propose a fast disparity estimation method considering temporal and spatial redundancy based on a dynamic programming for stereo matching. For the first step, the dynamic programming is performed to estimate disparity vectors with correlation between neighboring pixels in an image. Next, we efficiently compensate regions, which disparity vectors are not allocated, with neighboring disparity vectors assuming that disparity vectors in same object are quite similar. Moreover, in case of video sequence, we can decrease a complexity with temporal redundancy between neighboring frames. For performance comparison, we generate an intermediate-view image using the estimated disparity vector. Test results show that the proposed algorithm gives 0.8~2.4dB-increased PSNR(peak signal to noise ratio) compared to a conventional block matching algorithm, and the proposed algorithm also gives approximately 0.1dB-increased PSNR and 48~68%-lower complexity compared to the disparity estimation method based on general dynamic programming.

※ 본 연구는 한국과학재단 특정기초연구 (R01-2006-000-10199-0) 지원으로 수행되었음.

* 광운대학교 전자공학과 디지털 미디어 연구실 ({sneakes, redncap, sehwanpark}@kw.ac.kr, jsyoo@kw.ac.kr)

** 광운대학교 전자재료공학과 Digital Design & Test Lab. (dwkim@kw.ac.kr), *** 전자부품연구원(bhchoi@keti.re.kr)

논문번호 : KICS2008-04-194, 접수일자 : 2008년 4월 29일 최종논문접수일자 : 2008년 9월 17일

I. 서 론

현대 컴퓨터 산업의 발전과 멀티미디어 기술의 발달은 사용자들에게 더욱 사실적인 콘텐츠를 요구하게 되었고, 이러한 수요에 따라 대학, 기업 및 연구소에서는 실감 콘텐츠에 대한 연구를 활발하게 진행하고 있다. 그 중에서도 사용자로 하여금 입체감과 몰입감을 제공하는 3차원 입체 영상은 그 수요가 해마다 증가하고 있으며 각종 박람회나 전시회장의 이벤트 행사장에서 널리 사용되고 있다. 아울러 3DTV의 출시 및 보급을 위한 여러 기업과 정부기관의 노력은 3차원 입체 영상에 대한 부가가치와 미래 IT산업에서의 중요성이 어느 정도인지를 가늠케 한다.

3차원 입체 영상을 느끼게 하는 가장 일반적인 방법은 인간의 두 눈이 한 장면을 바라볼 때 생기는 양안 시차(binocular parallax)를 이용하는 스테레오 방식이다. 다시 말해 좌, 우 두 눈에서 본 것과 같은 영상을 여러 가지 방식을 통해 각 눈에 입력시켜 줌으로써 우리가 실생활에서 보는 것과 같은 깊이감을 느낄 수 있게 하는 것이다. 그러나 이러한 스테레오 방식은 시점이 고정되어 있기 때문에 여러 방향에서 관찰할 수 있는 자유도가 존재하지 않고, 무안경식의 경우에는 여러 사람의 동시 시청이 불가능한 단점이 존재하게 된다. 또한 스테레오 영상에서 더욱 확장된 다시점 영상(multi-view sequence)경우에도 데이터양의 증가와 시점간의 불연속성으로 인해 3차원 입체 영상구현에 많은 어려움이 따른다.

이러한 문제점을 보완하기 위해 두 시점의 영상으로부터 임의의 시점에 해당하는 영상을 생성하는 중간시점 영상 합성(IVR : intermediate view reconstruction)기법에 대한 연구가 활발히 이루어지고 있다. 중간시점 영상 합성 기술을 통해 다시점 영상에서의 효율적인 데이터 전송 및 수신이 가능해지며 시점간의 불연속 문제를 해결하게 된다. 또한 스테레오 영상에서 시점에 대한 자유도를 제공할 수 있게 된다. 중간시점 영상 합성에 있어 가장 중요한 과정은 좌, 우영상에 해당하는 두 시점의 영상이 주어졌을 때 영상 내 정합점(correspondence)을 찾아내는 변이 추정(disparity estimation) 방식이다. 변이 추정 방식은 정합 방법에 따라 크게 국부적 방법(local method)과 전역적 방법(global method)으로 나눌 수 있다¹⁾²⁾. 국부적 방법은 비교적 연산량이 적고 구현이 쉬운 반면에 가려진 영역(occlusion)이나 텍스처 정보가 부족한 영역(textureless)에서 오정합 발생 확

률이 매우 높다. 또한 정합창 사용으로 인한 경계 영역에서의 부정확한 변이가 할당되는 단점으로 인해 정확성 보다는 실시간이나 고속의 프로세스가 요구되는 분야에 사용이 되고 있다. 이러한 국부적 방법의 가장 대표적인 알고리즘은 한 영상의 작은 영역에 대하여 상관성이 가장 높은 영역이나 여러 값이 가장 낮은 영역을 참조 영상에서 탐색하는 블록 매칭(block matching) 방식이다. 또한 특징기반 매칭(feature-based matching)방식은 단순히 화소의 밝기 값을 이용하지 않고 모서리(edge), 구석(coner), 골(cave) 등 영상이 가지는 특징을 이용한다. 이외에도 화소의 밝기값 변화와 잡음 등에 취약한 블록 매칭 방식의 단점을 보완한 영상의 경사도(gradient method)를 이용한 방식이 있다.

이에 반해 전역적 방법은 텍스처 정보가 부족한 영역이나 가려진 영역에서의 오정합 발생을 줄이기 위해 비국부적 제약조건(nonlocal constraints)을 활용하는 방식이다. 이러한 전역적 방법은 정확한 변이를 추정할 수 있는 반면 계산 비용이 크다는 단점이 있으며 대표적인 방법으로는 그래프 컷(graph cut), 신뢰확산(belief propagation), 다이내믹 프로그래밍(dynamic programming)등의 알고리즘이 있다. 그래프 컷은 컴퓨터 네트워크 흐름제어에서 출발하였으며 최대-흐름/최소컷(max-flow/min-cut) 알고리즘을 이용하여 변이를 할당하는 방식으로 변이 추정에 가장 좋은 성능을 내는 방법 중 하나이다. 신뢰확산은 메시지 전달에 기반을 둔 반복 알고리즘으로 성능은 우수한 반면 완전히 수렴된 메시지를 얻기 위해서 많은 반복이 필요한 단점이 있다³⁾. 다이내믹 프로그래밍은 본래 어떠한 문제를 해결하는데 있어 최적화 문제에 대한 계산상의 복잡성을 줄이기 위한 수학적 방법이다. 다이내믹 프로그래밍 기반 변이 추정 방식이란 기준영상과 참조영상 사이의 관계를 이용하여 생성한 변이 공간 영상(DSI : disparity space image)으로부터 최적의 변이 경로를 찾아가는 방식을 말한다⁴⁾⁵⁾. 이러한 다이내믹 프로그래밍 기반 변이 추정 방식은 화소 단위의 정밀한 변이 맵을 얻을 수 있으며, 특징점 추출에 의한 오류를 막으면서 다른 전역적 방식들에 비해 수행 속도가 빠른 장점을 가지고 있다. 본 논문에서는 영상이 갖는 시·공간적 중복성을 이용한 다이내믹 프로그래밍 기반의 고속 변이 추정 방식을 제안한다. 영상 내 화소간의 높은 상관성을 이용하여 격행으로 변이 공간 영상을 생성 후 변이를 할당하며, 할당되지 않은 라인은 제안한 비용 함수를 이용하여 최종적으로 변이를 할당한다. 또한

동영상내에 존재하는 프레임간의 시간적 중복성을 이용하여 움직임이 존재하는 라인에 대해서만 변이 추정을 함으로써 수행시간을 단축시킬 수 있었다.

본 논문의 구성은 다음과 같다. 제 II장에서는 기존의 다이나믹 프로그래밍기반의 변이 추정 방식에 대하여 알아보고 제 III장에서는 제안된 변이 추정방식에 대하여 기술한다. 제 IV장에서는 성능 비교 및 분석을 위한 중간시점 영상 합성에 대하여 기술한다. 제 V장에서는 실험 및 결과를 분석하고 제 VI장의 결론을 끝으로 논문을 마친다.

II. 다이나믹 프로그래밍

그림 1은 기본적인 다이나믹 프로그래밍 기반의 변이 추정 기법을 나타낸다. 영상의 각 라인에 대하여 좌·우영상의 밝기값 차이를 나타내는 변이 공간 영상 생성 후, 적절한 비용(cost) 값을 이용하여 비용행렬을 생성한다. 마지막으로 최적의 경로를 탐색함으로써 변이를 할당하게 된다.



그림 1. 다이나믹 프로그래밍 기반의 변이 추정
Fig. 1. Disparity estimation by dynamic programming

2.1 변이 공간 영상(Disparity Space Image)

변이 공간 영상 생성은 다이나믹 프로그래밍 기반의 변이 추정 시 가장 먼저 수행되는 과정으로 이후의 비용행렬 생성과 최적경로 탐색에 영향을 주게 된다. 변이 공간 영상은 좌·우 영상의 n번째 스캔라인에 대한 정합 오류를 나타내는 것으로 그림 2에 좌·우영상의 100번째 스캔라인에 대해 생성된 변이 공간 영상을 나타내었다.

생성된 변이 공간 영상은 변이를 구하고자 하는 영상과 참조되는 영상의 n번째 스캔라인을 두 축으

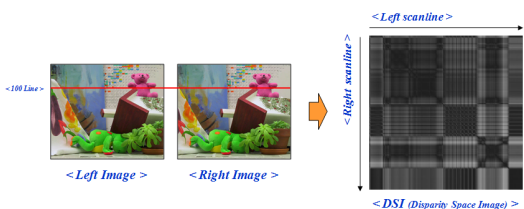


그림 2. 변이 공간 영상의 생성
Fig. 2. Generation of DSI(disparity space image)

로, 대응되는 좌표의 정합오류를 나타내며 영상내의 한 점은 식 (1)과 같이 나타낼 수 있다.

$$DSI_n(x_L, x_R) = \frac{1}{N} \sum_{i,j=-\frac{w}{2}}^{\frac{w}{2}} |I_L(x_L+i, n+j) - I_R(x_R+i, n+j)| \quad (1)$$

여기서 I_L 은 좌영상, I_R 은 우영상을 나타내며 w 는 정합창의 크기, N 은 정합시 사용된 블록내 화소 수를 나타낸다. x_L 과 x_R 은 각각 n번째 스캔라인에서 좌영상, 우영상의 좌표를 나타낸다.

식 (1)을 통하여 변이 공간 영상 내에 어두운 부분은 정합 오류가 작고 밝은 부분일수록 정합 오류가 크다는 것을 알 수 있다. 변이 공간 영상은 그 특징에 따라 다음과 같이 3개의 영역으로 나뉜다.

2.1.1 수직, 수평경계성분(Vertical, horizontal edge)

변이 공간 영상 내에 수직 경계성분은 좌영상의 경계성분에 해당되는 것으로서 경계 영역에서 화소값의 차이가 크다는 특성이 반영된 것이다. 마찬가지로 수평 경계 성분은 우영상의 경계성분에 해당되며. 이러한 수직, 수평 경계성분은 최적 경로를 탐색하는데 영향을 주게 된다.

2.1.2 질감 영역 성분 (Texture region)

좌영상과 우영상의 화소변화가 큰 영역이 서로 대응될 때 발생하는 영역으로 변이가 대응되는 좌표를 중심으로 화소의 변화가 크게 나타나기 때문에 변이를 쉽게 찾을 수 있는 영역이다.

2.1.3 평활 영역 성분 (Plain region)

평활 영역은 좌영상과 우영상의 화소변화가 거의 없고 평활한 영역이 서로 대응될 때 발생한다. 대응되는 변이를 중심으로 화소의 차이가 거의 존재하지 않으므로 변이를 찾는 데 어려움이 존재한다.

2.2 비용행렬(Cost Matrix) 생성

비용행렬은 앞서 얻어진 변이 공간 영상을 바탕으로 수직, 수평 경계성분을 고려한 비용값과 변이의 연속성을 고려한 비용값을 이용하여 생성하게 된다. 단순한 변이 공간 영상에서 최적의 경로를 찾게 될 경우 일반적인 화소기반의 변이 추정 방식과 크게 다르지 않게 되므로 비용행렬의 생성은 다이나믹 프로그래밍 기반의 변이 추정 방식에 특별성을 부여한다. 비용행렬 생성 시 고려되는 3가지 비용값은 식 (2)와 같다.

$$C_{plane}(x, y) = DSI(x, y) + Cost(x-1, y-1)$$

$$C_{horizontal}(x, y) = P_h + Cost(x-1, y) \quad (2)$$

$$C_{vertical}(x, y) = P_v + Cost(x, y-1)$$

여기서 C_{plane} 은 현재 화소의 변이가 이전 화소와 연속적인 변이값을 가질 때를 고려한 비용이며 $C_{horizontal}$ 과 $C_{vertical}$ 은 각각 변이 공간 영상 내에 수평 경계성분 및 수직 경계성분을 고려한 비용이다. P_h 와 P_v 는 수평, 수직 경계성분을 고려한 오프셋 값으로 가려진 영역을 판별하는데 영향을 미치게 된다. 위 3가지 값의 비교를 통한 최종적인 비용행렬 생성은 식 (3)과 같다.

$$Cost(x, y) = \min(C_{plane}, C_{horizontal}, C_{vertical}) \quad (3)$$

2.3 최적 경로 탐색(Optimal Path-Search)

최적 경로 탐색은 각각의 라인에 대해 생성된 비용행렬을 이용하여 최소 비용이 소요되는 경로를 찾는 과정이다. 목표점에서부터 시작점까지의 최소 비용 경로를 탐색하면서 최종적으로 변이를 할당하게 된다. 그림 3은 최적 경로 탐색을 통하여 변이가 할당되는 과정을 나타낸다.

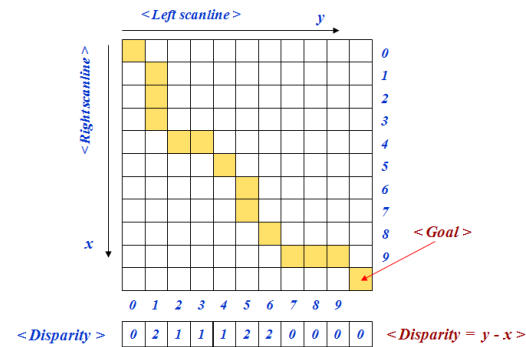


그림 3. 최적 경로 탐색 및 변이 할당
Fig. 3. Optimal path-search and disparity allocation

III. 제한하는 변이 추정 기법

3.1 공간적 중복성을 고려한 변이 추정

그림 4는 영상 내 공간적 중복성을 이용하기 위해 제안된 변이 추정기법이다. 일반적인 다이내믹 프로그래밍 기반의 변이 추정은 모든 라인에 대하여 변이 공간 영상 및 비용행렬을 생성하게 되는데 이

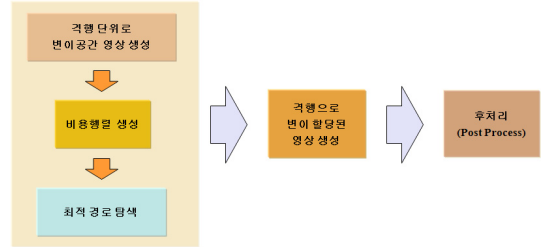


그림 4. 공간적 중복성을 이용한 변이 추정
Fig. 4. Disparity estimation considering spatial redundancy

로 인하여 많은 연산량이 필요하게 된다. 본 연구에서는 이러한 계산상의 복잡성을 줄이기 위하여 변이 공간영상 및 비용행렬을 격행 단위로 생성하고, 변이가 할당되지 않은 영역에 대해서는 후처리 과정을 이용하여 최종적으로 변이를 할당하게 된다.

일반적으로 영상 내 인접한 화소간에는 높은 상관성이 존재한다. 스테레오 정합에서 화소간의 상관성이 높은 영역, 즉 동일 영역의 변이 차는 매우 작거나 존재하지 않는다고 가정한다^[6]. 이러한 영상의 특성을 고려할 때 모든 라인에 대하여 계산 복잡성이 요구되는 다이내믹 프로그래밍 기반 변이 추정을 수행하는 대신, 격행으로(even or odd) 변이 추정 후 변이가 할당되지 않은 라인의 화소들을 주변 변이 벡터를 이용하여 할당하는 것은 복잡도(complexity) 측면에서 매우 효율적이다. 그림 5에 격행으로 변이가 할당된 영상을 나타내었다. 변이가 할당되지 않은 라인의 화소들은 이미 변이가 할당된 주변 영역으로부터 변이를 할당받게 되는데 이때 고려되는 주변 영역은 그림 6과 같다. 그림 6에서 A, B, C는 각각 변이를 할당하고자 하는 현재 화소의 위, 이전, 아래 화소이며 B 화소의 경우 영상의 맨 좌측 영역에서는 존재하지 않기 때문에 고려대상에서 제외된다. DV1, DV2, DV3는 각각 A, B, C 화소에 할당된 변이 벡터를 의미하며, 후처리 후에 현재 화소가 A 화소의 변이 벡터(DV1)로부터 할당된 것을 나타내

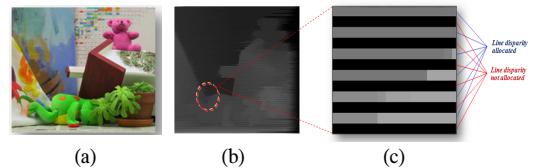


그림 5. 격행으로 변이가 할당된 영상
Fig. 5. Disparity-allocated image by interlace
(a) Original image (b) Disparity image allocated by interlace (c) Close-up of disparity image

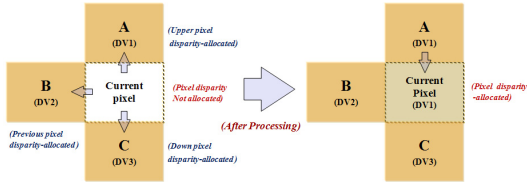


그림 6. 주변 변이 벡터로부터의 변이 할당
Fig. 6. Disparity allocation using neighboring disparity

었다. 이때 고려되는 주변 영역의 비용값을 구하기 위한 함수를 식 (4)에 나타내었다.

$$C(d_i) = T_{Neighbor(i)} + \lambda \cdot T_{Data}(d_i) \quad (4)$$

$$T_{Data}(d_i) = |I_R(x, y) - I_L(x + d_i, y)| \quad (5)$$

여기서 λ 는 T_{Data} 항에 대한 가중치이며 $T_{Neighbor}$ 항은 동일 영역으로 판단된 화소의 변이로 할당되어야 함을 고려한 것이다. T_{Data} 는 변이를 할당하기 위해 고려되는 주변 화소 변이 벡터들의 정확성을 판단하는 기준으로 사용된다. 만약 고려되는 변이 벡터가 정확하다면 현재 위치에서의 화소값과 참조 영상에서 정합되는 화소의 밝기값이 같거나 유사해야 한다는 것을 고려한 것이며 식 (5)와 같이 나타낼 수 있다. 여기서 $I_R(x,y)$ 은 현재 위치에서의 화소값을 나타내며, $I_L(x+d,y)$ 은 현재 위치에서 d 만큼 이동한 참조영상의 화소값을 나타낸다. 최종으로 현재 위치에서의 변이 할당은 식 (6)과 같이 나타낼 수 있다.

$$\hat{d}(x, y) = \min[C_{previous}, C_{Up}, C_{Down}] \quad (6)$$

여기서 $C_{Previous}$, C_{Up} , C_{Down} 은 현재 화소의 이전, 위, 아래의 비용값이며 최소값을 가지는 위치의 변이 벡터를 현재의 변이로 결정하게 된다. 일반적인 다이나믹 프로그래밍 기반의 변이 추정방식은 좌·우 영상의 동일 선상에 있는 정보들만을 이용하여 최종적인 변이를 할당하고 영상 내의 상하 상관성은 고려하지 않는다. 이로 인하여 잘못된 변이 정보가 수평방향으로 전파되는 단점이 존재하는데 제안한 방법의 경우 변이가 할당되지 않은 라인에 대하여 영상의 상·하 상관성을 고려하여 변이를 할당하기 때문에 수행속도의 개선뿐만 아니라 수평방향으로의 에러 전파에 대한 내성도 생기게 된다.

기존의 다이나믹 프로그래밍 기반의 변이 추정을 통해 얻은 변이 맵과 제안된 방법에 의해 얻은 변이 맵을 그림 7에 나타내었다. 그림 7의 (b)는 격행으로

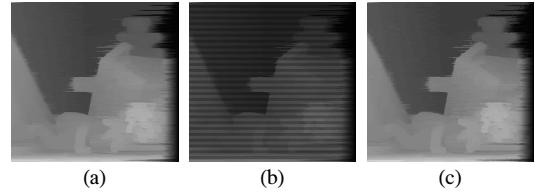


그림 7. 제안된 방법에 의한 변이 맵
(a) 기존의 다이나믹 프로그래밍 기반의 변이 맵
(b) 격행으로 할당된 변이 맵 (c) 후처리 후의 변이 맵

Fig. 7. Disparity image by proposed method
(a) Disparity image based on traditional dynamic programming (b) Disparity-allocated image by interlace (c) Disparity image with proposed post-process

변이가 할당되어 영상은 해상도가 반으로 줄기 때문에 전체적으로 어두워 보이는 것을 알 수 있다. 그림 7의 (c)는 제안된 방법에 의한 변이 맵으로 기존의 방법과 비교하여 가로방향의 에러 전파가 줄어든 것을 확인 할 수 있다.

3.2 시간적 중복성을 고려한 변이 추정

정지영상 내의 공간적 중복성 외에 초당 24, 혹은 30 프레임을 재생하는 동영상의 경우 앞, 뒤 프레임 간에 높은 중복성이 존재한다. 그림 8은 이러한 프레임간의 중복성을 이용하여 변이 추정 속도를 향상시키기 위해 제안한 방법이다.

먼저 현재 프레임과 이전 프레임의 차영상을 구하고 각 라인을 탐색한다. 탐색 과정을 통해 한 라인의 모든 화소값을 더한 후 임계값보다 작을 경우 이전에 구해진 변이 맵에서 해당하는 라인의 변이를 그대로 사용하게 되고 임계값보다 클 경우 움직임이 존재하는 라인으로 간주하여 다이나믹 프로그래밍 기반의 변이 추정을 수행한다. 동영상에서 배경 영역 부분은 움직임이 존재하는 하지 않는 경우가 대부분이기 때문에 위 과정을 통하여 수행시간을 크게 단축시킬 수 있게 된다. 그러나 현재 프레임과 이전

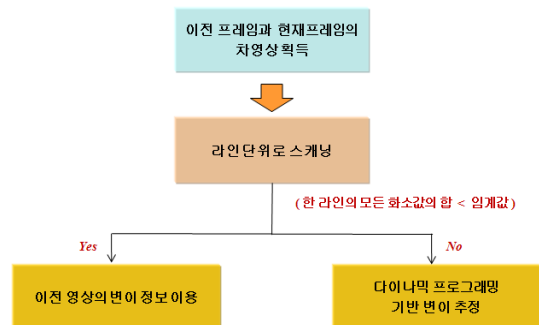


그림 8. 시간적 중복성을 이용한 변이 추정
Fig. 8. Disparity estimation considering temporal redundancy

프레임을 비교하여 아무 움직임이 존재하지 않는 라인이라 할지라도 각각의 카메라 특성과 잡음에 의해 미세한 화소값의 차이가 존재하게 된다. 본 논문에서는 이러한 차이 값을 가우시안 함수로 모델링하여 임계값을 결정하는데 이용하였다. 식 (7)는 차이 값 특성을 나타내기 위해 사용된 가우시안 함수이며, σ_x , σ_x^2 은 화소 차이값의 평균과 분산으로 각각 0과 2로 설정하였다.

$$f_x(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{(x-a_x)^2}{2\sigma_x^2}} \quad (7)$$

임계값은 각각의 화소 차이 값과 발생확률의 곱을 누적하여 구하며 식 (8)과 같이 나타낼 수 있다.

$$Threshold = \sum_{x=-5}^5 |x| f_x(x) H_{size} \quad (8)$$

여기서 x 는 발생하는 화소값의 차이를 나타낸다. H_{size} 는 영상의 가로 크기를 나타내며 각각의 차이 값 x 가 발생할 확률 $f_x(x)$ 와의 곱을 통하여 실제적으로 x 가 발생한 개수를 계산한다. 화소 차이 값의 최대, 최소 크기를 +5, -5로 정하였으며, -5부터 5까지의 누적 합을 통해 최종 임계값을 결정하게 된다. 화소값의 차이가 최대, 최소 크기를 넘어갈 경우 움직임이 존재하는 라인으로 판단하여 다이내믹 프로그래밍 기반의 변이 추정을 수행하게 된다.

그림 9는 이전 프레임과 현재 프레임의 차영상에서 각 라인의 누적 에러 값과 임계값의 비교를 통한 변이 할당을 나타낸다. 영상의 윗부분인 배경 영역에서는 움직임이 거의 존재하지 않기 때문에 누적 에러 값이 임계값보다 작게 된다. 이러한 영역의 라인은 이전 영상의 변이 맵으로부터 해당되는 라인의 변이를 할당받게 되고 영상 내 많은 움직임이 존재하는 객체 영역에 대해서는 다이내믹 프로그래밍 기반 변이 추정을 수행한다. 이러한 과정은 영상 내 많은 라인에 대하여 별다른 연산 없이 변이 할당을

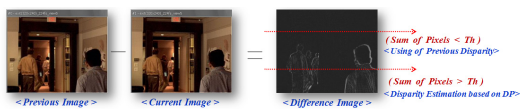


그림 9. 각 라인의 누적 에러 값과 임계값의 비교를 통한 변이 할당

Fig. 9. Disparity allocation by comparing accumulated-error with threshold value

가능하게 하며 동영상의 변이 추정시간을 단축하게 된다.

IV. 중간시점 영상 합성

(IVR : intermediate view reconstruction)

변이 추정 알고리즘 혹은 알고리즘 내의 파라미터에 따른 성능을 평가하기 위해서는 변이 맵의 객관적인 판단 기준이 필요하다. 변이 맵을 평가하기 위한 가장 일반적인 2가지 방법은 얻어진 변이 맵을 ground truth 데이터와 비교하여 통계적 에러를 구하는 방법과 변이 정보를 이용하여 새로운 시점의 영상을 생성 후 원래의 중간시점에 해당하는 영상과 비교하는 방법이다⁷⁾. 전자의 경우 텍스트처 정보가 부족한 영역, 가려진 영역을 제외한 모든 영역, 깊이 불연속 영역으로 영상을 분리한 후 각각의 영역에 대한 통계적 오류를 구하여 오류가 적을수록 성능이 우수한 것으로 판단하고⁸⁾, 후자의 경우 생성된 중간시점 영상과 본래 중간시점에 해당하는 카메라로부터 획득된 영상의 PSNR을 비교하여 성능을 판단한다. 본 논문에서는 중간시점 영상 생성을 통한 알고리즘의 성능평가를 실시하였으며 그림 10에 중간시점 영상 합성에 사용된 양방향 선형 보간법을 나타내었다⁹⁾.

그림 10에서 α 는 좌영상과 우영상의 거리를 1로 했을 때 중간시점 영상의 거리를 표현한 것이다. 생성하고자 하는 시점의 영상은 α 값이 1에 가까우면 우영상에 가까운 시점의 영상을, 0에 가까우면 좌영상에 가까운 시점의 영상을 나타낸다. 좌영상, 우영상, 중간시점 영상과 변이에 대한 관계는 식 (9)와 같이 표현할 수 있다.

$$I_\alpha(x+(1-\alpha)*d(x,y), y) = (1-\alpha)*I_{Left}(x+d(x,y), y) + \alpha*I_{Right}(x, y) \quad (9)$$

여기서 I_α 는 생성하고자 하는 중간시점 영상이며

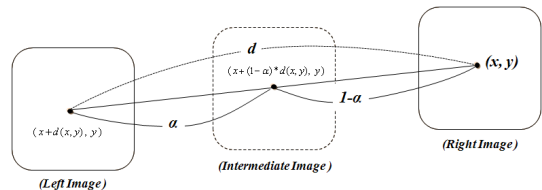


그림 10. 좌,우 영상으로부터의 중간시점 영상 생성

Fig. 10. Intermediate view reconstruction from a stereo image pair

$d(x,y)$ 는 변이 벡터를 나타낸다. I_{Left}, I_{Right} 는 각각 좌영상과 우영상을 나타내며 거리에 따른 가중치를 곱하여 더욱 자연스런 중간시점의 해당 화소를 표현한다. 그러나 변이 추정 시 가려진 영역으로 판단된 영역은 정확한 변이 정보가 할당되지 않기 때문에 중간시점 영상을 생성하는 과정에서 많은 문제를 발생시킨다. 그림 11은 가려진 영역이 발생하는 원인을 나타낸 것이며 가려진 영역은 2개의 경우로 나누어 처리하였다.

좌영상에만 존재하는 가려진 영역(left occluded-region)은 깊이가 작은 물체에 의해서 상대적으로 깊이가 큰 물체가 우영상에서는 존재하지 않는 영역이다. 발생한 지점의 좌우 변이 중 작은 쪽의 변이를 이용하여 식 (10)와 같이 좌영상에서 외삽(extrapolation)한다.

$$I_{\alpha}(x,y) = I_{Left}(x + \alpha * d, y) \quad (10)$$

여기서 $I_{Left}(x + \alpha * d, y)$ 는 현재 위치에서 변이 벡터 d 의 α 배 만큼 떨어진 좌영상의 화소값을 의미하며, $I_{\alpha}(x,y)$ 는 현재 위치에서의 생성된 중간시점 영상 화소값이다. 이와는 반대로 좌영상에는 존재하지 않고 우영상에만 존재하는 가려진 영역(right occluded-region)의 경우도 발생 지점의 좌우 변이 중 작은 쪽의 변이를 이용해 우영상에서 외삽하게 되며 식 (11)과 같이 나타낼 수 있다.

$$I_{\alpha}(x,y) = I_{Right}(x - (1 - \alpha) * d, y) \quad (11)$$

마찬가지로 $I_{Right}(x - (1 - \alpha) * d, y)$ 는 현재위치에서 변이 벡터 d 의 $(1 - \alpha)$ 배 만큼 떨어진 우영상의 화소값이며, $I_{\alpha}(x,y)$ 는 현재 위치에서의 생성된 중간시점 영상 화소값이다.

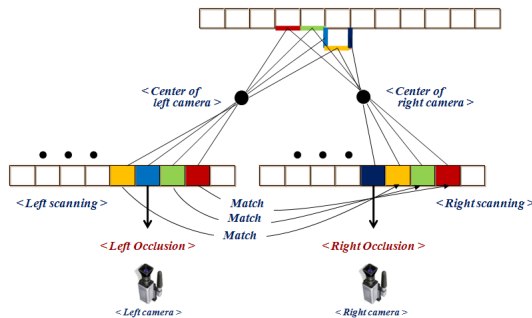


그림 11. 가려진 영역의 발생
Fig. 11. Generation of occluded-region

V. 실험 및 결과 분석

본 논문에서의 모의 실험은 Core 2 Duo CPU 2.2 GHz, 메모리 2GB의 사양에서 Visual C++로 구현되었으며 실험방법은 다음과 같다. 본 장의 제 1 절에서는 Middlebury 대학 컴퓨터 비전 연구실에서 제공하는 영상^[8]들을 이용하여 공간적 중복성을 이용한 변이 추정 기법의 성능을 기존의 알고리즘과 비교 분석한다. 제 2절에서는 MERL(Mitsubishi electric research laboratories)에서 제공하는 영상 시퀀스를 이용하여 시간적 중복성을 이용한 변이 추정 기법의 성능을 비교 분석한다.

5.1 공간적 중복성을 고려한 변이 추정 결과 및 분석

변이 추정 기법의 성능을 평가하기 위해 사용된 테스트 영상은 그림 12와 같다. 각 테스트 영상은 7대 혹은 9대의 카메라로부터 획득된 다시점 영상이며 시점 간에 보정(rectification) 과정이 수행되었다. 성능비교 방법은 Teddy 영상의 경우 2번과 6번, Moebius, Baby2, Doll 영상은 1번, 5번 카메라로부터 획득된 영상을 이용하여 중간시점의 영상을 생성하고 이를 실제 3번, 4번 카메라로부터 획득된 영상과의 PSNR을 구하고 각 알고리즘의 수행시간을 측정하였다. 성능 비교 알고리즘으로는 국부적 방법인 블록기반 정합 방식과 일반적인 다이나믹 프로그래밍 기반의 변이 추정 방식을 사용하였으며 표 1에 각 알고리즘에 대한 수행시간을 나타내었다.

표 1을 통해 제안한 방식이 기존의 다이나믹 프로그래밍 기반의 변이 추정 방식에 비해 약 48~49%의 수행속도가 단축됨을 알 수 있다. 또한 빠른 속도로 정합이 수행되는 8x8 블록기반의 변이 추정방

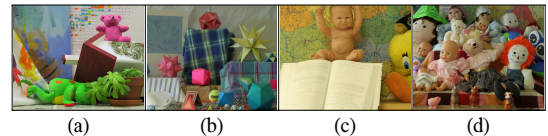


그림 12. 테스트 영상
(a) Teddy(450x375) (b) Moebius(450x375)
(c) Baby2(413x370) (d) Doll(427x370)
Fig. 12. Test sequences

표 1. 변이 추정 수행 시간 비교 (sec)

구분		Teddy	Moebius	Baby2	Doll
블록기반	8x8	0.30	0.31	0.23	0.32
	16x16	1.14	1.26	0.8	0.99
기존의 다이나믹 프로그래밍 기반		0.65	0.64	0.49	0.55
제안한 방식		0.33	0.32	0.25	0.32

표 2. PSNR 비교 (dB)

구분		teddy	moebius	baby2	doll
블록기반	8x8	27.1	26.9	28.4	25
	16x16	27.3	26.9	27.8	24.7
기존의 다이내믹 프로그래밍 기반		28	29.3	30	25.7
제안한 방식		28.1	29.3	30.1	25.8

식과의 비교에도 거의 같거나 약 10%의 계산량이 높은 것을 확인할 수 있었다. 이는 모든 라인에 대해서 수행되던 변이 공간 영상 및 비용행렬 생성, 최적경로 탐색 과정 횟수를 줄이고 후처리 과정을 통하여 효율적인 변이 할당이 이루어진 결과로 볼 수 있다. 표 2는 각 알고리즘으로부터 획득한 변이 맵을 이용하여 중간시점 영상을 생성 후 실제 해당 시점 영상과의 PSNR을 나타낸다.

본 논문에서 제안한 방식이 블록기반의 변이 추정 방식에 비해 0.8~2.4dB 높은 것을 확인할 수 있으며, 이는 일반적으로 전역적 방법이 국부적 방법에 비해 가려진 영역 혹은 텍스처 정보가 부족한 영역에서 더욱 정확한 정합을 수행한 결과로 볼 수 있다.

또한 기존의 다이내믹 프로그래밍 기반 변이 추정 방식과 비교에도 약 0.1dB 개선되었으며 이는 후처리 과정 시에 주변 영역을 고려하여 변이를 할당함으로써 일반적으로 발생할 수 있는 가로 방향의 에러 전파가 줄었기 때문이다.

결과적으로 기존의 전역적 방법인 다이내믹 변이 추정 알고리즘에 비하여 수행 속도 및 수행 결과 모두가 향상되는 것을 확인할 수 있었다. 또 다른 전역적 방법인 그래프 컷과 신뢰확산 알고리즘의 경우 다이내믹 프로그래밍 기반의 변이 추정 방식보다 좋



그림 13. 공간적 중복성을 고려한 변이 추정 결과
(a) 원 영상 (b) 변이 맵 (c) 실제의 중간시점 영상 (d) 생성된 중간시점 영상

Fig. 13. Result of disparity estimation considering spatial redundancy
(a) Original images (b) Disparity images (c) Original intermediate view images (d) Generated intermediate view images by proposed algorithm

은 성능을 나타내는 것으로 알려져 있으나 수행시간이 수십 초에서 몇 분의 시간을 필요로 하기 때문에 고속 변이 추정을 하는 본 알고리즘과의 비교 대상에서 제외하였다. 그림 13에 제안된 알고리즘을 이용하여 얻은 변이 맵과 생성된 중간시점 영상을 나타내었다. 그림 13의 (b)에서 검은 영역은 가려진 영역으로 판별된 부분이며 객관적 성능 도출을 위해 미디언 필터링(median filtering)과 같은 변이 맵 개선과정을 수행하지 않았다. 그림 13의 (c)는 각각의 좌, 우 영상의 중간에 해당하는 시점의 실제 영상이며 그림 13의 (d)는 그림 13의 (b) 정보를 이용하여 생성한 중간시점 영상이다. 그림 13 (c)의 실제 영상과 비교하여 볼 때 가려진 영역에 해당하는 가장 우측 경계영역을 제외하고 정교한 영상이 생성됨을 확인할 수 있다.

5.2 시간적 중복성을 고려한 변이 추정 결과 및 분석
현재와 이전 프레임간의 상관성을 이용한 변이 추정 기법의 결과 분석은 MERL에서 제공하는 Ballroom, Exit 영상 시퀀스를 사용하였으며, 그 특징은 표 3과 같다. MERL 외에도 여러 연구기관 및 대학에서 다양한 다시점 영상 시퀀스를 제공하고 있으나, 서로 대응되는 점이 같은 라인 선상에 존재한다는 에피폴라 제약조건을 사용하는 다이내믹 프로그래밍 특성상 보정 과정을 거친 MERL 시퀀스를 사용하였다. 또한 실시간에 근접한 변이 추정을 위해 영상을 320x240으로 다운 샘플링하여 실험에 사용하였으며, 표 4에 각 방법에 따른 변이 추정 시간을 비교하였다. 비교 방법은 일반적인 다이내믹 프로그래밍 기반의 변이 추정방식과 공간적 중복성을 이용한 변이 추정 방식, 시·공간 중복성을 이용한 변이 추정 방식을 사용하여 200 프레임의 변이 맵을 추정하는데 소요되는 시간을 측정하였다. 변위 탐색범위는 Exit 영상의 경우 60, Ballroom 영상의 경우 50으로

표 3. 테스트 영상 시퀀스의 특징

Data set	Sequence	Image Property	Camera Arrangement
MERL	Ballroom	640x480, 25fps (rectified)	8camera with 20cm spacing : 1D parallel
	Exit	640x480, 25fps (rectified)	8camera with 20cm spacing : 1D parallel

표 4. 변이 추정시간 비교 <200 frames 기준>

Method	일반적인 다이내믹 프로그래밍 기반	공간적 중복성을 이용한 변이 추정	시·공간 중복성을 이용한 변이 추정
EXIT	31.6sec	16.2sec	11.4sec
Ballroom	31.2sec	15.7sec	10.6sec

설정하였다. 실험결과 Exit 영상의 경우 일반적 다이내믹 프로그래밍 기반의 변이 추정 방식에 비해 공간적 중복성을 이용한 변이 추정을 경우 48%, 시·공간 중복성을 이용한 변이 추정을 할 경우 65%의 수행 속도가 개선되는 것을 확인하였다. Ballroom 영상은 공간적 중복성과 시·공간 중복성을 이용한 경우 각각 52%, 68%정도의 수행속도가 개선되는 것을 확인하였다. 그러나 MERL에서 제공하는 영상 시퀀스의 경우 카메라 간의 간격이 매우 크기 때문에 연속된 시점의 카메라가 아닌 1시점 떨어진 카메라를 이용하여 좌·우 영상을 획득할 경우 가려진 영역이 영상의 1/3이상을 차지하기 때문에 중간시점을 생성하기는 매우 어렵다. 따라서 생성된 영상의 PSNR 측정 평가는 따로 수행하지 않았다. 결과적으로 320x240 사이즈의 영상에서 초당 20 프레임에 가까운 변이 맵을 생성할 수 있었으며 시·공간 중복성을 고려한 변이 추정방식으로 얻어진 변이 맵을 그림 14에 나타내었다.

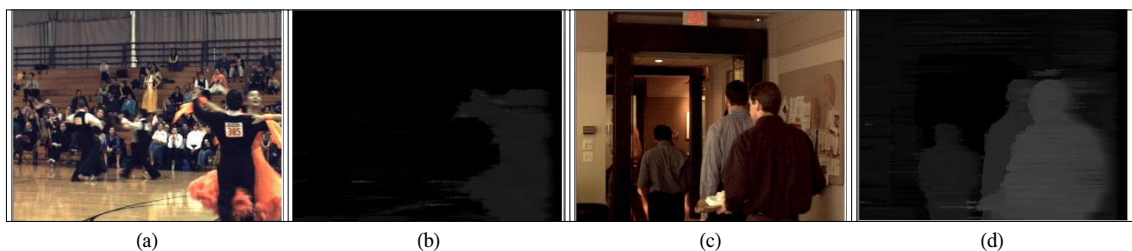


그림 14. 시·공간적 중복성을 고려한 변이 추정 결과
(a) Ballroom 영상 (b) Ballroom 영상의 변이 맵 (c) Exit 영상 (d) Exit 영상의 변이 맵
Fig. 14. Result of disparity estimation considering spatial and temporal redundancies

VI. 결 론

본 논문에서는 영상의 시·공간 중복성을 이용한 다이내믹 프로그래밍 기반의 변이 추정기법을 제안하였다. 영상 내에 존재하는 화소 간의 상관성과 동영상에서의 프레임 상관성을 이용하여 고속 변이 추정을 가능하게 하였으며 주변 영역을 고려한 후처리 작업으로 기존의 다이내믹 프로그래밍 기반의 변이 추정 방식보다 높은 PSNR 결과를 얻었다. 일반적으로 고속 변이 추정의 경우 연산량과 수행속도를 고려하여 국부적 방법의 변이 추정기법을 사용하는데 텍스처 정보가 부족한 영역이나 가려진 영역에서 많은 오정합이 발생하게 된다. 본 논문에서 제안한 방법은 이러한 단점을 보완하고 정교한 변이 추정을 하기 위해 전역적 방법을 이용하였으며 초당 20 프레임 정도의 변이 맵을 얻을 수 있었다. 또한 SIMD (single instruction multiple data) 기반의 MMX(multimedia extension)나 하드웨어로 알고리즘 구현 시 더욱 큰 사이즈의 영상에서 빠르고 정교한 변이 맵의 추출이 가능할 것이며 3D 모바일과 각종 단말기, 로봇 비전 분야로의 응용이 가능할 것이다. 그러나 제안한 변이 추정 방식은 에피폴라 제약조건을 가정함으로써 보정되지 않은 테스트 영상에 대해서는 보정된 영상에 비해 상대적으로 많은 오정합이 발생하는 문제점이 발생하였고, 상·하 상관성을 이용한 후처리 과정에도 불구하고 라인 단위의 변이 추정방식으로 인한 가로방향으로의 에러가 전파되는 문제점이 남아있다. 차후에는 영상 보정과정 연구를 통한 다양한 영상 시퀀스로의 적용과 에러 전파현상을 최대한 줄이면서 신뢰도 높은 변이 맵을 실시간으로 추출할 수 있는 변이 추정방식에 대한 연구가 필요하다.

참 고 문 헌

[1] Myron Z. Brown, Darius Burschka and Gregory D. Hager, "Advances in Computational Stereo," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.25, No.8, pp.993-1008, August 2003.

[2] Scharstein, D. and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *IJCV* 47(1), pp.7-42, 2002.

[3] 김봉겸, 임재권, "신뢰확산 알고리즘을 이용한 선 그룹화 기반 스테레오 정합," *전자공학회 논문지*

제42권 SP 제3호, 2005.

[4] P.H.S Torr and A. Criminisi, "Dense Stereo Using Pivoted Dynamic Programming," *Microsoft Research*, MSR-TR-2002-51, 2002.

[5] Ingemar J. Cox L. Hingorani, Satish B.Rao "A Maximum Likelihood Stereo Algorithm," *CVIU*, 63:3, pp.542-567, 1996.

[6] Zitnick, C. L., S. B. Kang, M. Uyttendaele, S.Winder and R. Szeliski, "High-quality video view interpolation using a layered representation," *Proceedings of SIGGRAPH* (ACM Transactions on Graphics), pp.600-608, 2004.

[7] R_Szeliski. "Prediction error as a quality metric for motion and stereo," *ICCV*, pp.781-788, 1999.

[8] <http://vision.middlebury.edu/stereo>

[9] 김미현, 강문기, 이철희, 최윤식, 손광훈, "경계-보존 방향성 평활화를 이용한 양안 영상의 변이 추정과 중간시점 영상의 재구성," *방송공학회논문지*, 제4권 제1호 pp.59-67, 1999.

윤 정 환 (Jung-hwan Yun)

준회원



2007년 2월 광운대학교 전자공학과 졸업
2007년 3월~현재 광운대학교 전자공학과 석사과정
<관심분야> Stereo Matching, 동영상 코덱, 3차원 영상처리

배 병 규 (Byung-kyu Bae)

준회원



2007년 2월 광운대학교 전자공학과 졸업
2007년 3월~현재 광운대학교 전자공학과 석사과정
<관심분야> 영상 압축, 3DAV

박 세 환 (Se-hwan Park)

준회원



2007년 2월 광운대학교 전자공학과 졸업
2007년 3월~현재 광운대학교 전자공학과 석사과정
<관심분야> 2D/3D 변환, 3차원 영상처리

송 혁 (Hyok Song)

정회원



1999년 광운대학교 제어계측공학과 졸업
2001년 광운대학교 전자공학과 석사
2006년 KETI 선임연구원
2006년 광운대학교 전자공학과 박사과정

<관심분야> 3D 영상 신호처리, 영상 압축

유 지 상 (Ji-sang Yoo)

정회원



1985년 2월 서울대학교 전자 공학과 졸업(공학사)
1987년 2월 서울대학교 대학원 전자공학과 졸업(공학석사)
1993년 5월 Purdue 대학교 전기공학과 졸업(Ph.D.)
1993년 9월~1994년 8월 현대전자 산업(주) 산전연구소 선임연구원

1994년 9월~1997년 8월 한림대학교 전자공학과 조교수
1997년 9월~2001년 8월 광운대학교 전자공학과 조교수
2001년 9월~2006년 8월 광운대학교 전자공학과 부교수
2006년 9월~현재 광운대학교 전자공학과 정교수
<관심분야> 웨이블릿 기반 영상처리, 영상압축, 영상인식, 비선형 신호처리

김 동 욱 (Dong-wook Kim)

종신회원



1998년 2월 한양대학교 전자공학과 졸업(공학사)
1985년 2월 한양대학교 대학원 졸업(공학석사)
1991년 9월 Georgia공과대학 전기공학과 졸업(공학박사)
1992년 3월~현재 광운대학교 전자재료공학과 정교수, 광운대학교 신기술 연구소 연구원

2000년 3월 ~ 2001년 12월 인티스닷컴(주) 연구원

<관심분야> 디지털 VLSI Testability, VLSI CAD, DSP 설계, Wireless Communication