

# 비선형 매니폴드 학습을 이용한 얼굴 이미지 합성

## (Face Image Synthesis using Nonlinear Manifold Learning)

조 은 옥 <sup>†</sup>   김 대 진 <sup>\*\*</sup>   방 승 양 <sup>\*\*\*</sup>  
(Eunok Cho)   (Daijin Kim)   (Sungyang Bang)

**요 약** 얼굴 구성 요소 각각에 대한 파라미터로부터 특정한 포즈나 표정을 갖는 얼굴 이미지를 합성하는 방법을 제안한다. 이러한 파라미터화는 얼굴 이미지의 표현과 저장, 전송을 효과적으로 수행할 수 있도록 한다. 그러나 얼굴 이미지의 변화는 고차원의 이미지 공간에서 복잡한 비선형 매니폴드를 구성하기 때문에 파라미터화 하는 것이 쉽지 않다. 이러한 문제점을 해결하기 위해, 얼굴 이미지에 대한 표현방법으로 LLE (Locally Linear Embedding) 알고리즘을 사용한다. LLE 알고리즘은 얼굴 이미지들 사이의 관계를 유지하면서 저차원의 특징 공간으로 투사된 매니폴드를 더욱 부드럽고 연속적으로 만들어준다. 그 다음, 특징공간에서 특정한 포즈나 표정 파라미터에 해당하는 포인트를 추정하기 위해 snake 모델을 적용한다. 마지막으로, 추정된 특징 값의 주변에 있는 여러 장의 얼굴 이미지들의 가중치 평균을 구해 합성된 결과 이미지를 만든다. 실험결과를 통해 제안된 방법을 이용하면 겹침 현상이 적고 포즈나 표정에 대한 파라미터의 변화와 일치하는 이미지를 합성한다는 것을 보인다.

**키워드** : 비선형 매니폴드 학습, Locally linear embedding, Snake 모델, 얼굴 이미지 합성

**Abstract** This paper proposes to synthesize facial images from a few parameters for the pose and the expression of their constituent components. This parameterization makes the representation, storage, and transmission of face images effective. But it is difficult to parameterize facial images because variations of face images show a complicated nonlinear manifold in high-dimensional data space. To tackle this problem, we use an LLE (Locally Linear Embedding) technique for a good representation of face images, where the relationship among face images is preserving well and the projected manifold into the reduced feature space becomes smoother and more continuous. Next, we apply a snake model to estimate face feature values in the reduced feature space that corresponds to a specific pose and/or expression parameter. Finally, a synthetic face image is obtained from an interpolation of several neighboring face images in the vicinity of the estimated feature value. Experimental results show that the proposed method shows a negligible overlapping effect and creates an accurate and consistent synthetic face images with respect to changes of pose and/or expression parameters.

**Key words** : Nonlinear manifold learning, Locally linear embedding, Snake model, Face image synthesis

### 1. 서 론

얼굴 이미지의 외양은 포즈나 표정에 따라서 변한다. 예를 들어, 대부분의 포즈는 수평방향과 수직방향에 대

한 변화로 나타낼 수 있다. 또한, 입과 눈 각각의 변화에 대한 조합으로 특정한 얼굴 표정을 나타낼 수 있다. 얼굴 이미지 자체는 고차원의 공간에서 표현되고 그들의 변화는 이러한 고차원의 데이터 공간에서 복잡한 비선형 매니폴드를 구성한다[1].

파라미터화는 데이터를 표현, 저장 그리고 전송할 때 유용하기 때문에, 이미지 자체보다는 포즈나 표정에 대한 적은 수의 파라미터로 얼굴 이미지를 표현하는 것이 좋다. 그러나 얼굴 이미지들은 고차원 공간에서 비선형 매니폴드를 구성하기 때문에, 얼굴의 외양과 포즈/표정

<sup>†</sup> 비 회 원 : 포항공과대학교 컴퓨터공학과  
pearleo@postech.ac.kr

<sup>\*\*</sup> 비 회 원 : 포항공과대학교 컴퓨터공학과 교수  
dkim@postech.ac.kr

<sup>\*\*\*</sup> 총신회원 : 포항공과대학교 컴퓨터공학과 교수  
sybang@postech.ac.kr

논문접수 : 2003년 2월 19일

심사완료 : 2003년 11월 4일

에 대한 파라미터 사이의 관계가 잘 부합되지 않는다, 그래서 이웃하는 이미지들 사이의 관계를 유지시키는 다른 데이터 공간으로 변환할 필요가 있다.

일반적으로 저차원 표현을 찾는 방법으로 주성분 해석(Principal component analysis)[2]과 다차원 척도법(Multi-dimensional scaling)[3]이 사용된다. 그러나 이들은 전체적으로 하나의 선형 변환을 기초로 하기 때문에 얼굴 이미지의 외관과 파라미터 사이에 불일치가 발생한다.

근래에, 비선형 매니폴드를 표현할 수 있는 비선형 차원 축소 방법인 locally linear embedding(LLE)[4,5]과 isometric feature mapping(Isomap)[6]이 제안되었다. LLE는 고차원 입력에 대해 이웃들과의 관계를 유지시키는 저차원의 표현을 찾는 무감독 학습 기법이다. Isomap은 데이터들이 구성하는 매니폴드의 표면을 따라 각 데이터들 간의 거리를 계산함으로써 본질적인 구조를 측정하는 방법을 사용한다. 더욱이, 이 두 방법들은 고차원 데이터를 저차원인 하나의 통합된 좌표계의 한 점으로 대응시킨다.

LLE와 Isomap은 비선형 차원 축소에 대한 비슷한 문제들에 성공적으로 적용된다. 그러나 Isomap의 결과는 각 데이터들 간에 모두 짝을 지어 그들간의 최단 거리를 유지할 수 있도록 최적화 되므로, 전체적으로 적용되는 제약조건과 모든 짝들 간의 거리를 이용한다. 그러나 LLE는 국소적으로 균형을 분석하기 위한 선형 계수를 사용하므로 거대한 동적 프로그래밍 문제를 해결할 필요가 없다. 매우 성긴 행렬을 축적시키려 하지만 이러한 구조는 시간과 메모리를 절약하기 위해 이용될 수 있다. 또한, LLE는 조각들로 나눌 때 겹쳐지는 부분에 있는 데이터들이 전체적인 구조에 대한 정보를 제공할 수 있다는 매니폴드 학습의 일반적인 원리를 보여 준다 [4]. 얼굴 이미지들이 이미지 공간에서 복잡한 비선형 매니폴드를 구성하기 때문에, 이러한 비선형 매니폴드 기법이 선형 차원 축소 방법보다 더 적당하다고 할 수 있다.

파라미터를 측정하는 일은 어렵기 때문에 특징 공간에 있는 점들에 대해 파라미터를 직접 명시할 수 없다. 그래서 파라미터 값이 쉽게 할당될 수 있는 가장자리의 특징 점들로부터 특정한 파라미터에 해당하는 특징 점을 추정한다. 이를 위해, 선이나 가장자리 또는 윤곽과 같은 이미지의 두드러진 특징을 찾아내는데 사용되는 snake 모델[7,8]을 사용한다. snake의 점들은 초기값이 윤곽에 근접하면 윤곽선을 향해 이동하는데 snake를 구성하는 이웃한 점들끼리는 동일한 거리를 유지한다. 그러므로 snake의 점들이 부드럽고 연속적인 매니폴드위에 위치하게 되서 snake의 점들이 각 파라미터에 대응

되는 특징점이 된다.

특정 파라미터 값에 대응되는 특징점을 구한다면, 추정된 특징점의 이웃에 해당하는 이미지들의 평균으로 합성 이미지를 만들 수 있다. 여기서 평균에 사용되는 가중치는 제약조건이 있는 최소사승법[5]을 이용해서 구한다.

예제 이미지들을 이용해서 특정한 파라미터에 대응되는 새로운 이미지를 생성하는 방법은 이미 제안되었다 [9]. 이미지들 사이의 광류추정 값으로 합성 그물망을 구성하므로, 3차원 모델링을 거치지 않아 시간과 공간의 절약할 수 있고 또 사진과 같이 실제 얼굴과 비슷한 이미지를 생성할 수 있다는 장점을 갖고 있다. 하지만 좋은 품질의 이미지를 얻기 위해서는 적지 않은 이미지를 사용해야 하는데, 모든 이웃하는 이미지들 사이의 광류를 추정해야 한다. 광류추정은 각 화소간에 이루어지는 연산이기 때문에 이미지의 크기에 따라 연산속도가 느려진다. 그리고 예제 이미지 각각에 대해 사람이 직접 파라미터를 부여해야 하기 때문에 부정확할 가능성이 많다는 단점을 갖고 있다.

이에 반해 본 논문에서 제시된 방법은 비선형 차원 축소 방법으로 데이터들을 저차원으로 변환해서 특징공간에서 연산이 이루어지고, 극단적인 이미지들에만 직접 파라미터를 부여해주면 되기 때문에 쉽고도 정확하게 파라미터화 할 수 있다.

비선형 축소법을 이용해서 각 이미지를 대표하는 저차원의 특징점을 구해서 간단한 입모양 이미지를 합성하는 것이 선형방법을 사용하는 것 보다 더 정확하다는 결과가 있다[10]. 본 논문에서는 비선형 축소법 중 데이터를 간단하고 정확하게 한 좌표계의 특징점으로 대응시키는 LLE 알고리즘을 사용하였고, 복잡한 표정과 포즈에 대응되는 파라미터를 구하기 위해 snake 모델을 적용하였다.

본 논문은 다음과 같이 구성된다. 2절은 LLE와 snake 모델에 대한 이론적인 배경을 설명하고, 3절에서는 이러한 기법들을 이용한 얼굴 이미지 합성방법을 기술한다. 4절에서는 실험을 통해 제안된 매니폴드를 기반으로 한 합성방법이 주성분 해석을 기반으로 한 방법보다 겹침정도와 파라미터와 외양간에 부합되는 정도에 대해, 더 정확하고 파라미터에 일치되는 얼굴 이미지를 합성 한다는 것을 보인다.

## 2. 이론적 배경

### 2.1 LLE 기법

LLE 기법은 비선형 표면을 굴곡을 거의 갖지 않는 작은 조각으로 나눈다. 각각의 작은 조각들은 편평해지고 조각들 간의 관계를 유지할 수 있도록 저차원의 공

간에서 재구성된다[11].

매니폴드로부터 얻어진 데이터가  $D$  차원의  $N$ 개의 실수 값으로 구성된 벡터  $\vec{X}_i$ 로 구성된다 가정하자. 각 데이터는 그와 이웃하는 데이터들로 복원될 수 있는 선형계수에 의해 조각들의 구조적인 특성을 나타낼 수 있다. 복원 오차는 다음과 같은 수식에 의해 측정된다.

$$\epsilon(W) = \sum_i |\vec{X}_i - \sum_j W_{ij} \vec{X}_j|^2 \quad (1)$$

가중치  $W_{ij}$ 는  $i$ 번째 데이터를 복원할 때  $j$ 번째 데이터의 기여도를 의미한다. 위의 복원 오차를 최소화하는 데에 두 가지 제약조건이 필요하다. 하나는 데이터  $\vec{X}_i$ 는 이웃하는 데이터들로부터 복원되므로  $\vec{X}_i$ 가  $\vec{X}_j$ 의 이웃에 속하지 않으면  $W_{ij}=0$ 이다. 두 번째는  $\sum_j W_{ij}=0$ 이 되도록 가중치  $W_{ij}$ 를 정규화(normalize)한다. 위의 두 제약조건에 대해 최적화된 가중치  $W_{ij}$ 는 최소자승 문제를 해결함으로써 얻어진다.

데이터가 낮은  $d(\ll D)$  차원의 부드러운 비선형 매니폴드 위에 위치한다고 가정하자. 직관적으로 이웃하는 데이터들의 고차원 좌표계를 매니폴드 위에 있는 하나의 포괄적인 좌표계로 대응시킬 수 있는 선형 대응관계가 존재하고 복원 가중치  $W_{ij}$ 는 변하지 않는다. 그래서 고차원의 이미지 공간에서 사용된 복원 가중치  $W_{ij}$ 는 저차원의 매니폴드 공간에서도 똑같이 사용될 수 있다.

LLE는 이웃하는 데이터들과의 관계를 유지시키도록 각 고차원의 데이터  $\vec{X}_i$ 가 매니폴드의 좌표를 나타내는 차원 데이터  $\vec{Y}_i$ 로 대응관계를 구성한다. 이는 다음의 함수를 최소화시키는  $d$ 차원의 좌표값  $\vec{Y}_i$ 를 선택함으로써 계산된다.

$$\phi(Y) = \sum_i |\vec{Y}_i - \sum_j W_{ij} \vec{Y}_j|^2 \quad (2)$$

위의 함수 또한 축소된 데이터 공간에서 국소적인 선형 복원오차를 기초로 하지만, 고정된 가중치  $W_{ij}$ 에 대해 최적화된 벡터  $\vec{Y}_i$ 를 결정한다. 최적의  $\vec{Y}_i$ 는  $d$ 개의 0이 아닌 가장 작은 고유치에 해당하는  $\phi(Y)$ 를 최소화하는 고유벡터들의 정렬된 집합이다[4].

**2.2 SNAKE: Active Contour Model**

snake는 외부에서 제한하는 힘에 의해 유도되고, 가장자리와 같은 특징들로 끌어당기는 이미지 자체의 힘에 영향 받아 에너지를 최소화하는 곡선을 그리는 점들의 집합이다[7]. 에너지 최소화는 단계별로 수행되고 각 단계에서의 움직임은 현재 단계의 에너지가 최소가 됨을 의미한다. 이를 실행하기 위해 에너지는 세 개의 에너지 항목, 연속성(Continuity), 곡률(Curvature), 이미

지 자체의 에너지로 정의된다.

$$E_{snake} = \int_0^1 (\alpha(s)E_{cont} + \beta(s)E_{curv} + \gamma E_{\text{energy}}) ds. \quad (3)$$

$$E_{cont} = \frac{|V_x(s)|^2}{2}, \quad (4)$$

$$E_{curv} = \frac{|V_{ss}(s)|^2}{2}, \quad (5)$$

$$E_{\text{energy}} = w_1(I(x, y) - w_{edge}|\nabla I(x, y)|)^2 + w_{term}(\frac{\partial \theta}{\partial n_{\perp}}) \quad (6)$$

$V(s) = (x(s), y(s))$ 는 호의 길이에 대해 윤곽선의 좌표를 나타내는 함수이고,  $V_x(s)$ 와  $V_{ss}(s)$ 는 각각  $s$ 에 대한  $V(s)$ 의 1차도함수와 2차도함수이다.

첫 번째 항목인 연속성은 snake의 인접한 두개의 점이 등거리를 갖도록 한다. 두 번째 항목인 곡률은 호의 길이에 대한 외부의 각도 변화를 측정한다. 다른 힘이 영향을 끼치지 않는다면 곡률은 snake를 불룩한 원형의 모양을 형성하게 한다. 세 번째는 가장 중요한 항목인 이미지 자체의 에너지이다. 이 에너지는 기울기의 크기로부터 얻어진다. 파라미터  $\alpha(s)$ ,  $\beta(s)$ ,  $\gamma$ 는 세 항목들의 상대적인 효과의 균형을 맞추기 위해 사용된다[12].

**3. LLE 기반의 이미지 합성**

제한된 얼굴 이미지 합성 방법은 다음과 같이 세단계, LLE 알고리즘을 이용한 특징공간 구성과 특징점 추정, 얼굴 이미지 합성으로 구성된다.

**3.1 LLE 알고리즘을 이용한 특징공간 구성**

LLE 알고리즘의 성능은 데이터의 개수에 의해 크게 영향을 받는다[13]. 만약 데이터가 충분히 존재하지 않는다면, 부분적인 선형을 보장할 수 없다.  $N$ 개의  $D$  차원의 이미지들을  $\vec{X}_i, i=1, \dots, N$ 이라고 하자.  $\vec{X}_i$ 와 대응하는  $d$ 차원의 특징점  $\vec{Y}_i$ 는 최소자승 문제와 Rayleitz-Ritz 이론[14]을 이용해서 계산할 수 있다. 그래서 모든 데이터 이미지들에 해당하는 특징점들  $\vec{Y}_i, i=1, \dots, N$ 을 구할 수 있다. 이 특징점들은 일반적으로 부드럽고 연속적인 매니폴드 공간을 구성한다.

**3.2 특징점 추정**

모든 가능한 포즈나 표정을 갖는 얼굴 이미지를 얻기도 어렵고 두 극단적인 경우( $p=0$  and  $p=1$ )를 제외하고는 각 특징점에 파라미터 값을 할당하는 것도 어렵다. 이러한 한계점을 극복하기 위해, 특정한 포즈나 표정의 파라미터 값에 대응하는 특징점을 추정하며 이를 위해 snake 모델을 사용한다.

snake의 수렴을 보장하기 위해서는 초기점이 매니폴

드에 가깝도록 설정되어 있어야 한다. 그러므로  $\bar{Y}$ 의 부드러운 매니폴드 곡선을 한 축을 기준으로  $M$ 등분하고 이 점들을  $M$ 개의 점들로 구성된 snake의 초기값으로 해서 수식 (3)의 함수를 최소화함으로서 snake  $\bar{S}_i, i=1, \dots, M$ 을 구할 수 있다.

snake 모델은 snake에 존재하는 인접한 점들 사이에 등거리를 유지하므로, snake를 구성하는 각 점  $\bar{S}_i$ 는 정확하게 각 파라미터에 해당하는 특징점이 된다. 그림 1은 파라미터 값이  $p=0$ 과  $p=1$ 인 두개의 특징점들로부터 포즈 파라미터  $p=0.25$ 와  $p=0.5, p=0.75$ 에 대응되는 특징점에 대한 추정결과를 보여준다.

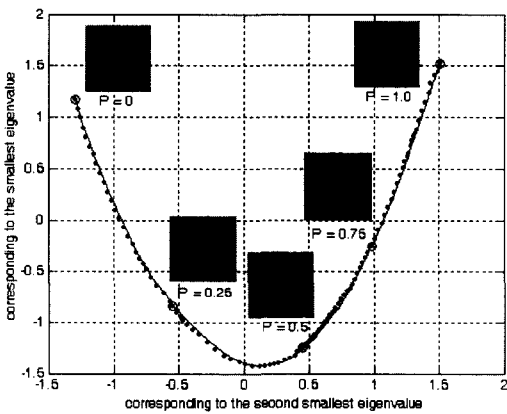


그림 1 snake 모델을 이용한 포즈 추정

### 3.3 얼굴 이미지 합성

위의 두 단계를 통해 얻어진 snake를 구성하는 점  $\bar{S}_i$ 에 해당하는 얼굴 이미지를 합성해야 하는데 다음과 같이 두 가지 방법을 사용할 수 있다. 한 방법은  $\bar{S}_i$ 와 가장 가까운 이미지  $I_{MN}$ 을 선택하는 것이다.

$$I_{MN} = X(Y_{MN}(S(p))) \quad (7)$$

$p$ 는 파라미터 값이고,  $S(p)$ 는 파라미터  $p$ 에 해당하는 snake를 구성하는 점,  $Y_{MN}$ 은  $S(p)$ 와 가장 가까운 특징점  $\bar{Y}_i, 1 \leq i \leq N$ ,  $X$ 는 특징점  $\bar{Y}_i$ 에 해당하는 이미지를 의미한다.

다른 방법은 추정된 특징점과 가까운  $K$ 개의 이웃 이미지들의 가중평균 이미지를 결과이미지로 사용하는 것이다.

$$I_{KNV} = \sum_{k=1}^K W_k X_k(Y_k(X(p))) \quad (8)$$

가중치  $W_k$ 는 최소자승 문제를 풀어서 얻을 수 있고,  $Y_k$ 는 snake를 이용해서 추정된 특징점  $S(p)$ 와 가까운  $K$ 개의 이웃하는 데이터들 중  $k$ 번째 특징점,  $X_k$ 는 특

징점  $Y_k$ 에 해당하는  $k$ 번째 얼굴 이미지이다. LLE 알고리즘이 각 데이터를 여러 개의 이웃하는 데이터들로부터 선형적으로 근사치를 구할 수 있는 부분적인 선형성을 가정하기 때문에, 이 두번째 방법이 타당하다고 할 수 있다.

## 4. 실험 결과 및 고찰

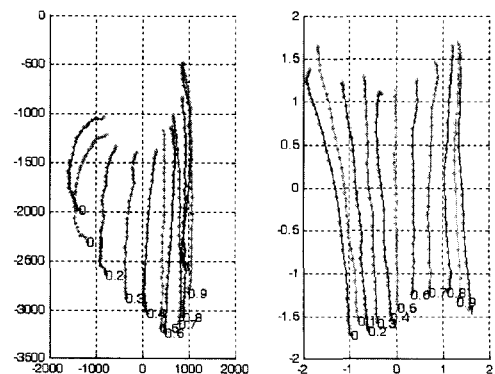
정확한 합성 이미지를 만들기 위해서는 거대한 양의 예제 이미지를 필요로 하기 때문에 정확하게 모든 포즈를 갖는 얼굴 이미지를 합성하는 것은 어렵다. 더욱이 각기 다른 표정과 포즈를 갖는 얼굴 이미지를 합성하는 것은, 각 포즈에 대해 모든 가능한 표정의 예제 이미지가 필요하므로 더욱 어렵다. 그래서 포즈와 표정에 대한 이미지 합성 문제를 나누어서 다루었다.

### 4.1 포즈 이미지 합성

포즈는 수평방향과 수직방향으로 변화할 수 있다. 예제 포즈 이미지들은 한 사람이 고개를 각 11개의 수평 위치에서 수직방향으로 움직이는 비디오 연속물로부터 얻어진다. 학습에 사용되는 데이터는 379개의  $64 \times 64$  흑백 이미지로 구성된다. 이는 그림 2와 같이 각 수평 위치에서 고개의 수직적인 움직임을 나타내는 11개의 곡선으로 이루어진 매니폴드를 구성한다.

차원축소 방법으로 주성분해석과 LLE 알고리즘을 비교한다. 그림 2는 2차원의 특징점들이 구성하는 두개의 다른 방법에 의한 매니폴드를 보여준다. 그림 2(a)에서 보듯이 주성분해석에 의해 얻어진 매니폴드의 일부분이, 파라미터  $[0, 1]$ 과  $[0.8, 1]$ , 겹쳐있다. LLE에 의해 구성된 매니폴드가 주성분 해석에 의해 구성된 매니폴드보다 더 부드럽고 규칙적임을 알 수 있다.

본 실험에서, 수평방향 포즈 파라미터 값에 해당하는 특징점들은 데이터 이미지들의 포즈가 많은 부분에서



(a) 주성분해석에 의한 매니폴드 (b) LLE에 의한 매니폴드

그림 2 두개의 다른 매니폴드 비교

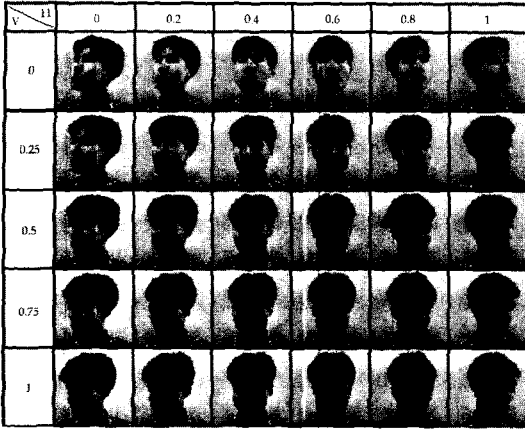


그림 3 수평방향과 수직방향에 대한 포즈 파라미터에 따른 합성된 얼굴 이미지

연속적이지 않으므로 정확하지 않다. 그러므로  $K$ 개의 이웃하는 이미지들의 가중평균을 이용하는 경우 다른 포즈를 갖는 얼굴 이미지들 간의 겹침현상을 발견할 수 있다. 또한 가장 가까운 이미지를 선택하는 경우에는, 다른 포즈 파라미터 값을 입력했음에도 불구하고 같은 얼굴 이미지를 결과로 내기도 한다. 그러나 대부분의 경우에 주어진 포즈 파라미터와 일치하는 포즈 이미지를 합성할 수 있다. 그림 3은 합성된 결과 얼굴 이미지들을 보여준다. 여기서 'H'는 수평방향, 'V'는 수직방향을 의미한다.

주성분해석과 LLE, 두개의 매니폴드 구성 방법을 이용해서 합성된 포즈 이미지의 품질을 비교하도록 한다. 공정한 비교를 위해 차원축소 단계를 제외한 특징점 추정과 가장 가까운 이미지를 선택하는 합성방법을 모두 동일하게 사용한다. 그림 4는 고정된 수평 포즈 파라미터 1.0과 수직 포즈 파라미터  $[0, 0.13]$ 에서 0.01의 간격으로 변화하는 파라미터에 대한 결과이미지를 보여준다. 그림 4(a)는 수직 포즈 파라미터와 이미지 사이의 많은 불일치와 수평 포즈 파라미터가 고정되어 있음에도 불구하고 포즈의 수평방향 움직임을 보여준다. 그러나 그림 4(b)는 수평 방향으로 아무런 포즈 변화도 없고 수직 포즈 파라미터와 합성된 포즈 이미지들이 일치함을 보여준다.

특정 공간의 매니폴드에서 각 곡선에 대해 101개의 이미지를 합성하는 경우 전체적으로 1,111개의 결과 이미지를 얻을 수 있다. LLE를 사용하는 경우 26장의 포즈 불일치 이미지를 얻을 수 있고 주성분 해석을 이용하는 경우 78장의 포즈 불일치 이미지를 얻을 수 있다.

이미지 합성 방법으로 추정된 특징점과 이웃하는  $K$ 개 이미지들의 가중평균을 사용할 때 합성된 포즈 이미

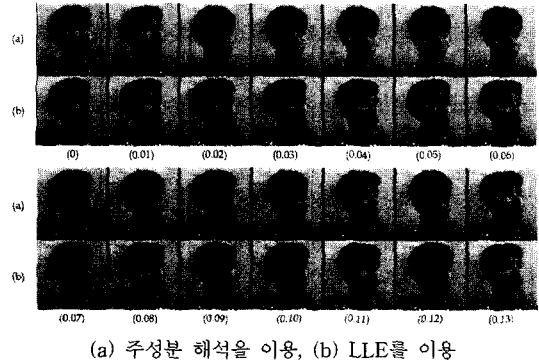


그림 4 고정된 수평 포즈 파라미터와 변화하는 수직 포즈 파라미터에 대응되는 결과이미지의 예

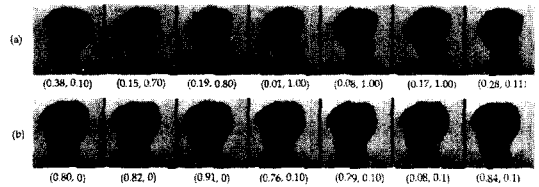


그림 5 두개의 다른 매니폴드, (a) 주성분 해석, (b) LLE, 에 의해 얻어진 결과이미지에서 겹침 정도 비교

지들의 겹침 정도를 비교하겠다. 추정된 특징점에 근접한 4개의 이웃하는 이미지들로부터 결과 이미지를 합성하므로 겹침현상은 피할 수 없다. 그림 5는 두개의 다른 매니폴드 구성 방법에 따른 결과 이미지를 보여준다. (a)와 (b)는 각각 주성분 해석과 LLE에 해당한다. 주성분 해석에 의한 매니폴드로부터 얻어진 결과 이미지들은 LLE에 의한 매니폴드로부터 얻어진 결과 이미지들에 비해 더 많고 심한 겹침현상을 갖는다. 이는 주성분 해석을 이용한 매니폴드가 LLE 매니폴드에 비해 겹쳐 있는 부분이 많고 연속성을 갖지 못하기 때문이다.

#### 4.2 표정 이미지 합성

얼굴 표정은 입과 눈, 눈썹, 볼 등과 같은 얼굴 구성 요소들의 변화로서 다루어질 수 있다. 여기서는 입과 눈의 변화(열림과 닫힘)만을 고려하였다. 눈의 변화는 얼굴의 다른 부분에 별로 영향을 끼치지 않으므로 눈 부분은 개별적으로 다룰 수 있다. 입에 대해 합성된 얼굴 이미지에 눈 합성 이미지를 붙여서 표정을 생성함으로써 포즈에 대한 2차원의 파라미터를 사용하는 것에 비해 훨씬 적은 예제 이미지를 사용해서 더 좋은 결과를 얻을 수 있다.

입 모양의 합성을 위한 예제 이미지는 한 사람이 입을 다문 상태에서 천천히 열리는 비디오 연속물로부터

추출한다. 학습에 사용되는 데이터는 128×128 크기의 43장의 흑백이미지로 구성된다. 눈 모양의 합성에 사용되는 예제 이미지도 입과 마찬가지로 눈을 감은 상태에서 점점 크게 뜨는 비디오 연속물로부터 이미지를 추출한다. 두 눈과 눈썹을 포함하는 부분 이미지만을 선택한다. 눈 이미지 데이터는 128×128 크기의 이미지로부터 선택된 68×32 크기의 53장의 이미지로 구성된다.

그림 6은 눈과 입의 벌림 정도에 따른 파라미터 값이 0에서부터 1까지 0.25의 간격으로 변화할 때의 합성된 표정 이미지이다. 입과 눈의 벌림 정도에 대한 파라미터 값을 따라 합성된 표정 이미지들이 자연스럽게 정확하게 변화하고 겹침현상도 나타나지 않는다.

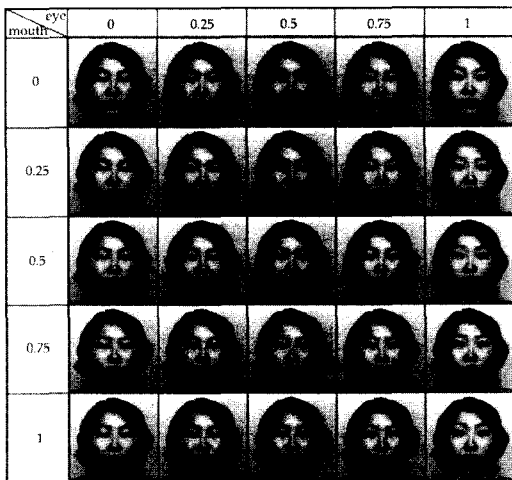


그림 6 각각의 눈과 입의 벌림 정도에 대한 파라미터에 따른 합성된 표정 이미지

### 5. 결론

LLE 알고리즘과 같은 비선형 매니폴드 학습과 snake 모델을 이용한 특징점 추정을 이용해서 특정한 포즈나 표정을 갖는 얼굴 이미지를 합성하는 방법을 제안하였다. 포즈는 수평방향과 수직방향을 따라 2차원의 파라미터로 나타낼 수 있고, 표정 이미지는 눈과 입의 벌림 정도를 파라미터화해서 나타낼 수 있다.

제안된 얼굴 이미지 합성 방법을 사용하면 적은 양의 얼굴 이미지와 파라미터만 저장하고 있으면, 주어진 파라미터 사이에 존재하는 임의의 파라미터에 대응되는 포즈와 표정을 갖는 얼굴 이미지를 생성할 수 있으므로 적은 저장 공간을 필요로 하기 때문에 유용하다.

얼굴 이미지 합성은 매우 낮은 차원의 특징공간(본 실험에서는 2차원의 특징 공간이 선택되었다.)에서 수행되므로 시간적인 면에서도 효과적이라고 할 수 있다.

제안된 얼굴 이미지 합성 방법은 LLE 매니폴드가 파라미터 값의 변화에 일치하는 부드러운 모양을 형성하고, 기초를 이루는 데이터 구조를 잘 표현하기 때문에 더욱 정확하다고 할 수 있다. LLE 알고리즘은 특징공간에서 이웃하는 데이터들과의 관계를 유지시키도록 하기 때문에 상대적으로 적은 겹침 현상을 보인다. snake 모델은 인접하는 데이터들끼리 등거리를 보장하므로 파라미터의 값과 특징점들이 일치하도록 만든다.

### 참고 문헌

- [1] S. Gong, S. J. McKenna, A. Psarrou, "Dynamic Vision(From Images to Face Recognition)," Imperial College Press, 2000.
- [2] R. O. Duda, P. E. Hart, D. G. Stork, "Pattern Classification," A Wiley-Interscience Publication, 2001.
- [3] T. Cox and M. Cox, "Multidimensional scaling," Chapman & Hall, 1994.
- [4] S. T. Roweis, L. K. Saul, Nonlinear Dimensionality Reduction by Locally Linear Embedding, Science, Vol. 290 22, pp. 2323~2326, 2000.
- [5] L. K. Saul, S. T. Roweis, "Think Globally, Fit Locally: Unsupervised Learning of Nonlinear Manifolds," University of Pennsylvania Technical Reports, MS-CIS-02-18.
- [6] J. B. Tenenbaum, V. Silva, J. C. Langford, "A Global Geometric Framework for Nonlinear Dimensionality Reduction," Science, Vol. 290 22, pp. 2319~2322, 2000.
- [7] M. Kass, A. P. Witkin, D. Terzopoulos, "Snakes: A Active Contour Models," International Journal of Computer Vision(1), pp. 321~331, 1998.
- [8] D. Willians, M. Shah, "A Fast Algorithm for Active Contours and Curvature Estimation," CVGIP: Image Understanding, pp. 14~26, 1992.
- [9] T. Ezzat, T. Poggio, "Facial Analysis and Synthesis Using Image-Based Models," Proceedings of the Second International Conference on Automatic Face and Gesture Recognition, October, pp. 116~121, 1996.
- [10] C. Bregler, S. M. Omohundro, "Nonlinear Image Interpolation using Manifold Learning," In Advances in Neural Information Processing Systems 7, 1995.
- [11] K. Patch, "Tools cut data down to size," Technology Research News, March 14, 2001.
- [12] R. Jain, R. Kasturi, B.G. Schunck, "Marchine Vision," New York: McBraw -Hill, 1995.
- [13] A. Hadid, O. Kouropteva, M. Pietikainen, "Unsupervised Learning using Locally Linear Embedding: Experiments with Face Pose Analysis," 16th International Conference on Pattern Recognition, 2002.

- [14] G. H. Golub, C. F. Van Loan, "Matrix Computations," The Johns Hopkins University Press, 1996.



조 은 옥

1997년 3월~2001년 2월 충남대학교 컴퓨터 공학과 학사(B.S.). 2001년 3월~2003년 2월 포항공과대학교 컴퓨터 공학과 석사(M.S.)



김 대 진

1981년 2월 연세대학교 전자공학과(학사). 1984년 2월 KAIST 전기 및 전자공학과(석사). 1991년 8월 Syracuse University, Electrical and Computer Eng. (박사). 1984년 3월~1986년 12월 한국방송공사 기술연구소. 1992년 3월~1999년 6월 동아대학교 컴퓨터공학과 부교수. 1999년 7월~현재 포항공과대학교 컴퓨터공학과 부교수. 주관심분야는 지능시스템, 멀티미디어 처리



방 승 양

1966년 일본 Kyoto대학 전기공학에서 학사. 1969년 서울대학교 전기공학에서 석사. 1974년 미국 University of Texas 전산학에서 박사를 받았음. 미국 Wayne State University, NCR, Bell 연구소 등에서 근무하다가 1981년 귀국. 한국전자기술연구소 시스템부 실장, 부장 역임, (주)유니온시스템 전무. 1986년부터 포항공대 컴퓨터공학과 교수. 현재 뇌연구센터 소장. 관심분야는 패턴인식, 신경회로망