

# 知的画像符号化에 있어서 表情分析和 合成

崔 昌石\*                      原島 博\*\*                      武部 幹\*  
 \* 日本金沢大学工学部                      \*\* 日本東京大学工学部

## Analysis and Synthesis of Facial Expressions in Knowledge-Based Image Coding

Chang Seok CHOI\*, Hiroshi HARASHIMA\*\* and Tsuyosi TAKEBE\*  
 \*Kanazawa university, Japan                      \*\*The university of Tokyo, Japan

New image coding system for facial images called 'Knowledge-based image coding' is described, in which input image is analyzed and output image is synthesized using analysis results. Analysis and synthesis method of facial expressions are presented. Synthesis rules are determined on the basis of facial muscles and are also used in analysis process to produce a faithful reconstruction of the original image. A number of examples are shown.

### 1. 서론

画像은, 그내용을 直觀的으로 파악할 수 있는 情報 伝達媒体이지만, 방대한 伝送帯域이 画像通信의 실현에 커다란 장애요인이 되고 있다. 이를 위해, 전송대역을 압축하는 高能率符号化에 대한 연구가 집중적으로 진행되고 있다.

현재, 연구의 주류를 이루는 高能率画像符号化는, 要素의 振幅情報에 주목하여, 時空間相關이나 振幅分布 등의 統計的知識을 이용하는 波形符号化가 대부분이다.

그러나, 카메라를 통해서 입력되는 대부분의 화상은 3次元世界的인 2次元投影이라 할 수 있고, 対象을 어느 정도 한정할 수 있는 경우, 対象物에 대한 특별한 「知識」을 이용하는 새로운 부호화방식을 생각할 수 있다. 예를들면, 대상물체에 대한 3次元構造 등이 얻어질 수 있다면, 3次元構造를 「知識」으로서 이용하는 부호화가 가능할 것이다. 구체적으로는 TV전화, TV회의 등에서는, 符号化對象을 얼굴에 제한할 수 있고, 입력 화상으로부터 얼굴의 3次元構造를 복원하여 「知識」으로서 이용하는 부호화방식이 생각된다<sup>[1-5]</sup>. 이와같은 부호화방식을 分析合成符号化(Analysis Synthesis Coding) 또는 知的符号化(Knowledge-Based Coding, Model-Based Coding, Intelligent Coding)라 부르고 있다.

이와같은 부호화에서는, 얼굴의 3차원모델을 「知識」으로 하여 입력화상을 분석하고, 그결과로서 構造情報, 運動情報, 表情情報 등을 추출한다. 또, 추출된 정보

로부터 表情을 합성하여 출력화상을 표시하는 것이 중심과제이다. 이때, 어떤 表情記述파라미터를 이용하여, 본래의 表情情報을 잃지 않고, 表情의 分析和 合成을 효율적으로 할 것인가가 문제가 된다. 예를들면, FACS(Facial Action Coding System)의 AU(Action Unit) [6]는 심리학분야에서 제안된 조직적인 表情記述法으로, 表情의 分析和 合成에 있어서도 表情기술의 규범이 될 수 있다. 知的画像符号化에서 FACS의 AU를 表情記述파라미터로서 이용하는 경우, 表情의 자동분석 방법 및 합성법을 확립할 필요가 있다.

본문에서는, FACS의 AU를 表情기술파라미터로서 도입하여, 얼굴의 3차원모델을 이용한 表情의 合成法과 分析法에 대해서 검토하고자 한다.

### II. 얼굴画像의 知的符号化<sup>[1-3]</sup>

#### 1. 기본개념

知的符号化는, 종래의 파형레벨의 부호화와는 달라, 부호화對象에 대한 「知識」을 적극적으로 이용해서, 삼으로 본질적인 정보만을 송수신하는 통신방식이다. 이 부호화방식의 기본구성은 그림 1과 같다.

送信側에서는 얼굴의 3차원모델을 이용하여 입력 화상으로부터 構造分析, 運動推定, 表情分析을 행한 결과로서 構造情報, 運動情報, 表情情報, 모델更新情報 등을 추출하여, 受信側에 전송한다.

受信側에서는 전송된 정보로부터 얼굴의 3차원모델

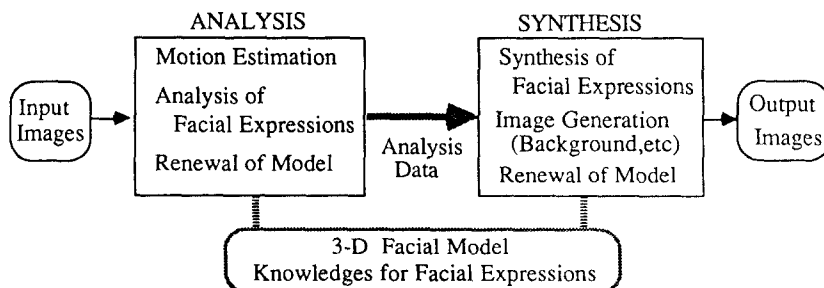


Fig.1 Knowledge-based coding system for facial images

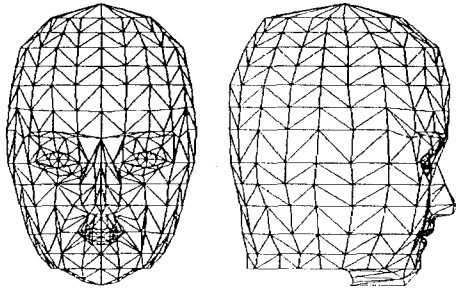


Fig.2 3-D wire frame general face model



Fig.3 A full-face neutral image and adjusted wire frame general face model

을 변형해서 표정을 합성한다. 이때, 필요에 따라 更新情報를 이용해서 보다 자연스런 표정을 합성한다. 또, 전송정보에 포함되어 있지않는 背景등은 로컬적인 圖像生成을 이용한다.

이와같이, 얼굴화상의 지적부호화에서는, 분석결과만을 전송하기때문에 대폭적인 情報壓縮이 가능하고, 분석결과와는 가공이 가능하기때문에 유연성이 풍부한 미래의 통신방식으로서 주목되고 있다.

### 2. 얼굴의 3 차원모델의 구성

표정을 분석하고, 그 결과로부터 자연스런 표정을 합성하기 위해서는, 대상인물의 얼굴에 대한 충실한 3 차원모델이 필요하다. 여기서는 마네킹의 얼굴을 작은 삼각형으로 근사한, 그림2와 같은 3 차원 와이어 프레임 모델(Wire Frame Model)을 얼굴의 一般形状모델로 했다. 이 와이어프레임 모델에 대상인물의 無表情正面像에 그림3과 같이 정합하여 개인 와이어프레임 모델을 얻는다. 이때, 얼굴의 크기와 깊이를 3 차원이 판변환(Affine Transform)을 이용하여 조정한다. 특히, 얼굴전체의 윤곽 및 각부의 (예를들면, 눈썹·눈·코·입·턱등)의 윤곽을 가능한 정확히 맞춘다. 이어서, 개인 와이어프레임 모델에 正面像의 輝度情報를 投影해서 대상인물의 얼굴에 대한 3 차원모델을 얻는다.

얼굴의 3 차원모델의 구성은, 2 차원의 얼굴화상으로부터 3次元構造를 복원하는 작업임으로, 일단 모델이 구성되면 임의의 시점으로부터 본 얼굴화상을 쉽게 합성할 수 있다.

### III. 표정분석과 합성에 있어서 기본사고방식

表情記述파라미터로서 FACS의 AU를 도입하여, 그것을 이용한 표정분석과 합성에 있어서 기본사고방식에 대

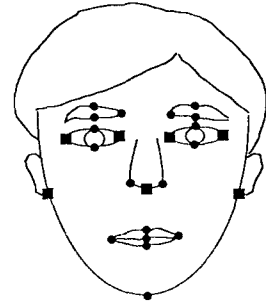


Fig.4 Feature points used in motion estimation and analysis/synthesis of facial expressions

해서 기술한다.

#### 1. FACS에 대하여<sup>[6]</sup>

FACS(Facial Action Coding System)란 심리학분야에서 제안되어 있는 表情記述法이다. 顔面筋肉을 해부학적으로 고려하여, 표정을 AU(Action Unit)라는 44개의 基本動作으로 분해하고 있다. FACS에 의하면, 모든 표정은 AU또는 그 조합으로 기술가능하다고 되어 있다.

그러나, FACS의 AU는 컴퓨터상에서 실현되어 있는 것이 아니며, 強度도 정량적으로 취급되고 있지않다. 만약, FACS의 AU가 정량화되어 컴퓨터상에 실현된다면, 얼굴화상의 自動分析·合成을 비롯하여 심리학·정신병리학등 다방면에 응용이 가능할 것이다.

#### 2. 기본적인 사고방식

(1)얼굴의 움직임은, 頭部의 大局인 運動과 表情變化에 따른 얼굴각부위의 局部的인 움직임으로 분리한다. 이는, 각각을 분리하여 정보를 구조적으로 기술하기위함이다. 이를 위해 그림4와 같이 特徴點을 정한다. 두부의 運動推定에는 점(■)를, 각부위의 움직임 검출에는 점(●)를 이용한다. 각각을 運動推定特徴點, 表情特徴點이라고 부르기로 한다.

(2)얼굴화상의 확대·축소 및 個人差에 대처하기 위해, 표정기술파라미터를 기준량으로 正規화한다. 기준량으로서 눈썹·눈·입등의 종횡폭(그림7,8)을 이용한다.

(3)표정분석·합성과정에서, 본래의 표정정보를 잃지않고 재현성을 좋게 하기위해, 분석과정에서 합성규칙을 이용한다.

(4)표정분석과 합성의 결과가 心理学知見으로도 적절하고, 가능하면 표정의 추상적인 의미의 추출 및 합성이 가능케 한다.

### IV. 表情合成法<sup>[4,5]</sup>

表情合成은, 無表情의 와이어프레임모델을 변형하여, 無表情像의 輝度值를 삼각형별로 텍스처 매핑(Texture Mapping)해서 행해진다. 이과정을 그림5에 나타낸다. 이때, 표정기술파라미터로서 FACS의 AU를 이용하는 경우, 각각의 AU에 대한 와이어프레임 모델의 變形規則을 정하지않으면 안된다. 표정은 안면근육의 수축·이완에 의해서 변화하는 것이기때문에, AU변형규칙을 정할 때, 안면근육을 참고로한다. 안면근육을 그림6에 나타낸다. 원내숫자는 AU의 번호이고, 숫자의 위치는 근육의 수축방향이다. AU의 내용은 표1과 같다.

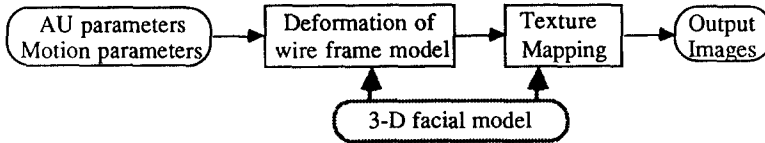


Fig.5 Synthesis procedure of facial expressions

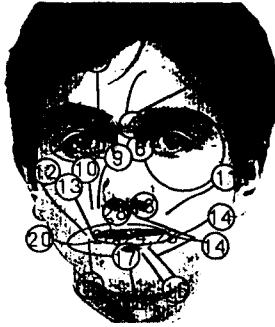


Fig.6 Locations and action directions of facial muscle fibers

1. 얼굴상부에 관련한 AU의 실현

눈에는, 내측을 올린다(AU1), 외측을 올린다(AU2), 내측을 하방비스듬이 끈다(AU4) 등이 있다. 각각의 AU와 그 조합의 실현례는 그림7과 같다. AU1의 경우는 눈의 内端点B1을, AU2의 경우는 중앙점B2를 올리고, AU4의 경우는 점B1을 하방비스듬이 끌어서, 특징점근방의 점은 직선으로 근사한다. 각각의 특징점은 AU의 강도에 비례해서 이동한다. AU1+2, AU1+2+4와 같은 AU 조합에 대한 特徴点의 移動量은 각각의 AU에 대한 특징점이동량의 線形和로 한다.

눈에는, 눈을 크게 뜬다(AU5), 눈을 감는다(AU43), 웅크린다(AU48) 등의 AU가 있다. 이러한 AU의 실현에는 눈중앙의 특징점E1, E2를 상하로 움직여서 2차곡선으로 근사한다.

2. 얼굴하부에 관련한 AU의 실현

얼굴하부의 AU는 입에 집중해 있고, 턱과 코에 관한 것도 있다. 입의 움직임은 수직·수평·경사적인 움직임, 꺾이진 움직임 및 기타로 분류할 수 있다.

대표적인예로서 아랫입술을 내린다(AU25), 입양단을상방비스듬이 올린다(AU12)와 그들조합의 실현례를 그림8에 나타낸다. AU25의 경우는 아래입술의 특징점M4, M5를 내리고, AU12의 경우는 입술양단점M1, M1'을 상방비스듬이 올려, 특징점근방은 2차곡선으로 근사한다.

3. 합성례

현재, 표1과 같이 34종류의 AU를 실현했다. 그중 일부의 AU와 AU조합에 대한 합성상을 그림9에 나타낸다. 표정은, AU합성규칙에 따라 좌상의 무표정을 변형하여 합성했다. 무표정에서 얻을 수 없는 주름(눈썹사이, 코등, 코부터 입양단사이)과 齒 등의 휘도정보는 原画像에서 갱신한 정보를 이용했다<sup>[5]</sup>.

V. 表情分析法

표정분석은 표정합성의 逆過程이다. 즉, 얼굴각부위의 변형된 형상으로부터 AU를 추출하는 과정이다. 얼굴화

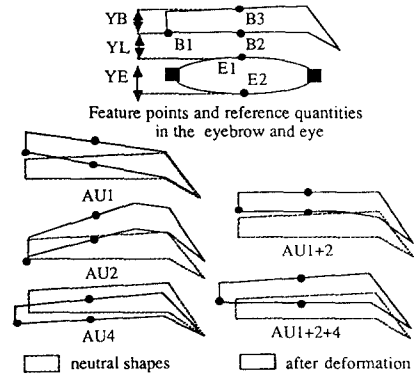


Fig.7 Deformation of eyebrow shapes for Action Units

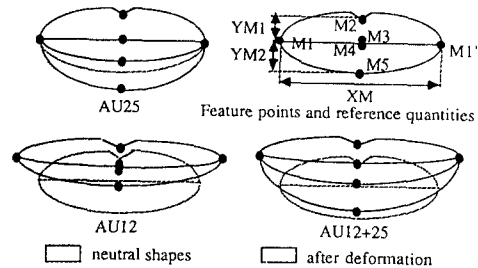


Fig.8 Deformation of lip shapes for Action Units

상으로부터 얻어진 특징점을 이용한 AU의 추출과정을 그림10에 나타낸다. 여기에서는, 얼굴화상으로부터 특징점은 추출되어 있다고 가정한다.

1. 頭部の 3次元運動 推定<sup>[3]</sup>

두부를 剛体로 간주하여, 운동파라미터를 2개의 얼굴화상으로부터 추정한다. 3차원공간에서 微小運動을 가정하면, 강체상의 점의 운동식은 식(1)과 같다.

$$P' = RP + T \quad (1a)$$

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & \phi & -\phi \\ -\phi & 1 & \theta \\ \phi & -\theta & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} \quad (1b)$$

P, P'는 운동전후점의 x, y, z 좌표를 요소로 하는 벡터(이하, 좌표벡터라 한다)이다. R은, 각축의 회전량이  $\phi, \theta, \psi$ 인 회전운동을 나타내는 행렬이다. T는, 각축의 평행이동량  $\Delta x, \Delta y, \Delta z$ 을 요소로 하는 벡터이다. 2차원화상으로부터 얻어지는 운동전후점의 좌표(x, y), (x', y')와 깊이정보로서 와이어프레임의 z 좌표를, 식(1b)에 대입하면, 식(2)가 얻어진다. 여기에서, 화상은 平行投影을 가정하여,  $\Delta z=0$ 로 한다.



Fig.9 Images synthesized from AU and AU combination parameters

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} 1 & \phi & -\phi \\ -\phi & 1 & \theta \\ \phi & -\theta & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \\ 0 \end{bmatrix} \quad (2)$$

식(2) 를 식(3) 과 같이 변형하여, 3 점 이상의 복수 점에 대해서 연립시켜, 自乘誤差를 최소로 하는 운동 파라미터( $\phi, \theta, \psi, \Delta x, \Delta y$ ) 를 구할 수가 있다. 실제로, 두 부의 운동추정에는 그림 4 의 7 점 (■) 을 이용했다.

$$\begin{bmatrix} X' - X \\ Y' - Y \end{bmatrix} = \begin{bmatrix} Y & 0 & -Z & 1 & 0 \\ -X & Z & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \phi \\ \theta \\ \psi \\ \Delta x \\ \Delta y \end{bmatrix} \quad (3)$$

2. 입력화상의 頭部位置의 修正  
두부의 운동과 각부위의 움직임을 분리할 목적으로, 입력화상과 무표정화상의 두부위치를 일치시킨다. 각 화상의 運動推定特徵点座標벡터를  $P'$  라하면, 식(1) 으로부터 근사적으로 식(4) 가 성립한다.

Table1 Realized Action Unit list

AU No	AU Name	AU No	AU Name
1	Inner Brow Raiser	17	Chin Raiser
2	Outer Brow Raiser	18	Lip pucker
4	Brow Lowerer	20	Lip Stretcher
5	Upper Lid Raiser	23	Lip Tightener
6	Cheek Raiser	24	Lip Pressor
7	Lid Tightener	25	Lips Parts
8	Lips Toward	26	Jaw Drop
9	Nose Wrinkler	27	Mouth Stretcher
10	Upper Lid Raiser	28	Lips Suck
11	Nasolabial Furrow Deepener	29	Jaw Thrust
		30	Jaw Sideways
12	Lip Corner Puller	32	Bite
13	Sharp Lip Puller (Cheek Puffer)	35	Cheek Suck
		41	Lip Droop
14	Dimpler	42	Slit
15	Lip Corner Depressor	43	Eyes Closed
		44	Squint
16	Lower Lip Depressor	45	Blink
		46	Wink
			Eyeballs Turn

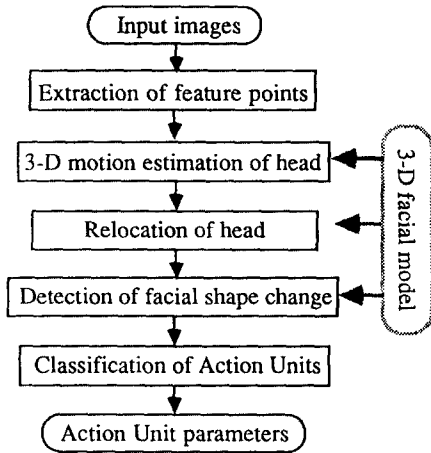


Fig.10 Analysis procedure of facial expressions

$$P = R^{-1} (P' - T) \quad (4)$$

입력화상의 表情特徴点座標벡터를  $P'$  라고하고, 식(4) 과 같이 변형하면, 식(5) 와 같다.

$$F = R^{-1} (F' - T) \quad (5)$$

$F$  가, 입력화상의 두부위치를 무표정화상과 일치시켰을 때, 입력화상의 표정특징점좌표벡터가 된다. 이때, 입력화상으로부터 얻어지지 않는 깊이정보는 와이어프레임의  $z$  좌표를 이용한다.

### 3. 表情变化의 特徴量算出

표정변화의 특징량은 입력화상과 무표정화상의 얼굴 각부위의 형상을 비교해서 산출한다. 즉, 무표정화상의 表情特徴点座標벡터를  $F_0$ 라 하면, 表情变化의 特徴量벡터  $D$ 는  $D=F-F_0$ 로 한다.

### 4. AU의 判別

변형된 얼굴 각부위의 형상은 AU조합의 형상으로 볼 수 있다. 이 형상으로부터 AU의 조합을 어떻게 추출할 것인가가 문제가 된다. 본문에서는, 하나의 어프로치로서 각각 AU에 대한 특징을 미리 조사해 놓고, 얼굴각부위의 형상변화로부터 특징에 일치한 AU를 차례로 추출하는 방법을 검토한다.

#### (1) 얼굴상부에 관련된 AU의 판별

눈썹의 경우, 그림 7의 AU1+2+4 조합을 예로서 AU의 판별방법을 설명한다. 먼저, 가장 特徴이 있는 AU4로부터 시작한다. AU4는 眉간을 좁게 하는 AU로서, 눈썹의 内端点B1이 중앙쪽으로 이동하게 된다. 이것은, 表情变化特徴量벡터  $D$ 의 要素DXB1(점B1의 X축방향의 변화량)의 正負(X축은 중앙쪽을 負方向, 바깥쪽을 正方向으로 한다)로부터 판단된다. AU4의 強度는 DBX1을 눈썹의 기준량(YB)로 정규화한 양으로 한다. 그림 7의 AU1+2+4의 형상으로부터 AU4 합성시에 부여하는 변화량을 제거하면, AU1+2의 형상이 된다. 다음에, AU1에 特徴이 있으므로 DYB1(점B1의 Y축방향의 변화량, Y축은 하방을 負方向, 상방을 正方向으로 한다)로부터 AU1를 판별하여 동일한 작업을 반복하면, AU2의 형상이 남게 된다. 최후로, DBY2로부터 AU2를 판별하면, 눈썹에 관한 표정변화특징량벡터  $D$ 는 0 벡터가 되고 AU판별은 종료한다. 이상을 木構造로 정리하면

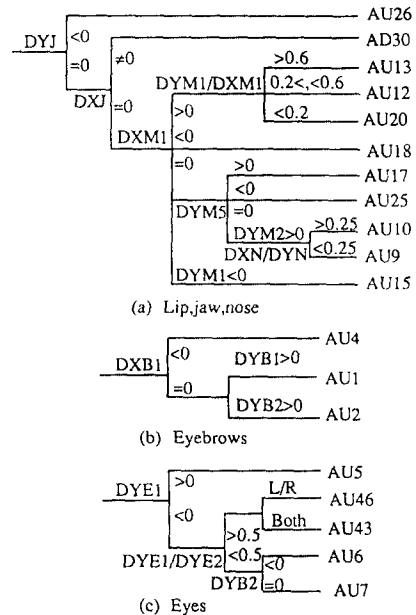


Fig.11 Tree structures for AU classification

그림11(b) 와 같다.

눈에 관련된 AU에는 上下瞼의 상하운동에 의하여 합성시에 소개한 것처럼 AU5, AU43, AU46 등이 있으나, AU판별목구조만을 그림11(c)에 게재한다.

#### (2) 얼굴하부에 관련된 AU의 판별

얼굴하부에는, 입에 관한 AU가 많고, 표현력이 풍부한 반면, AU판별에는 그만큼 복잡하다. AU판별에 있어서는, 형상변화가 선명한 입과 턱을 주로 이용하고, 코는 보조적으로 이용한다. 구체적으로 그림 8의 AU12+25를 예로서 AU판별과정을 설명한다. 가장 판별하기 쉬운 AU는 턱에 관한 것이다. 턱에 변화가 없으면 입양단에 특징이 있다. AU12는 입양단을 상방비스듬히 올리는 것으로, X, Y축방향의 이동량DXM1, DYM1이 正인 경우이다. AU12의 強度는, DXM1을 기준량(XM)으로 정규화한 양으로하고, AU12의 합성시 형상변화량을 AU12+25의 형상으로부터 제거하면, AU25의 형상이 남게된다. AU25는 아래입술을 내리는 AU로, DYM5가 負인 경우이다. 이와같은 방법으로, 얼굴하부에 대한 AU를 판별하는 목구조를 그림11(a)에 나타낸다. 여기서 이용한 점J, N은 그림4의 턱과 코양단에 정한 특징점이다. 여기서 주의해야 할 점은 판별순서에 의하여 결과가 다를 수 있다는 것이다. 예를들면, AU9의 경우는 입에 수된 특징이 있지만, 코, 눈, 눈썹까지도 변화한다. 이와같은 것을 고려해서, AU판별은 그림11과 같이 얼굴의 하부로부터 시작한다.

### 5. 분석례

표정분석결과와 정당성은, 그것으로부터 합성된 표정의 재현성에 의해서 평가될 것이다. 실제 얼굴화상에서 얻어진 특징점으로부터 표정을 분석하고, 분석결과를 이용하여 합성한 画像를 原画像과 그림12에 비교한다. 합성화상은, 화상의 무표정상을 변형하여 얻어진 것으로, 원화상의 표정을 비교적 양호하게 재현하고 있다는 것을 알수있다.

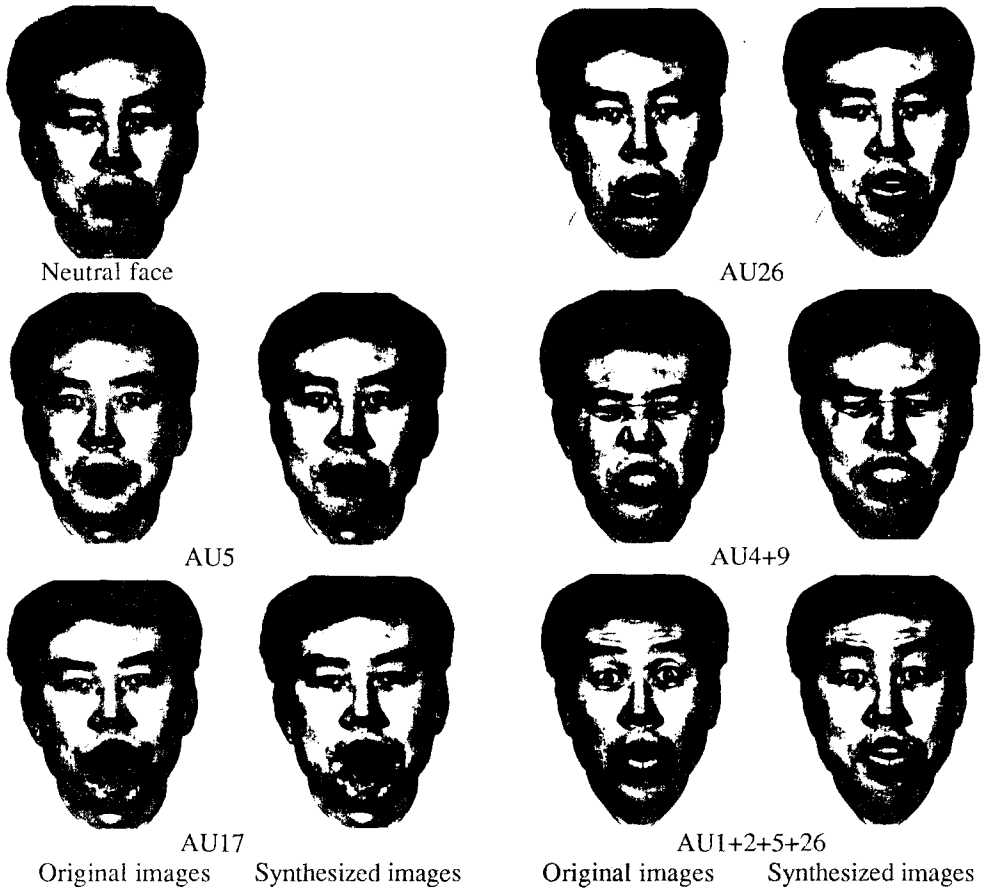


Fig.12 Comparison of original images with images synthesized from analysis results

## VI. 결 론

知的画像符号化方式을 개관하고, 중심이 되는 表情分析과 合成에 대해서 검토했다. 表情分析과 合成過程에서, 表情기술파라메티르로서 심리학분야에서 제안된 FACS의 AU를 도입하여, 기본사고방식에 대해서 서술했다. 表情合成에 있어서는 顔面筋肉을 고려하여, 34종류의 AU를 컴퓨터상에서 실현하였다. 이때, AU를 얼굴 각부위의 폭 길이등으로 정규화하였다. 합성례는 비교적 자연스런 표정이 일어났다.

表情分析에서는 頭部の 3 차원운동파라메티르를 추출하고, 두부의 大局的인 運動과 얼굴 각부위의 움직임을 분리하여, 분석결과를 구조적으로 기술했다. AU의 효율적인 판별을 위한 木構造를 설계했으며, 表情의 再現性을 위해 분석과정에서 합성규칙을 이용했다. 분석결과를 이용해서 합성한 표정은 원화상의 표정을 비교적 양호하게 재현하였다.

본문에서 제안된 표정분석법과 합성법은 그결과로부터 지적화상부호화에 유효하다는 것을 알 수 있다.

표정분석과정에서는 특징점추출을 가정하였으나, 지적화상부호화의 실현에 있어서는 대단히 중요한 과정이다. 금후, 특징점추출에 대해서도 검토할 예정이다.

## 참 고 문 헌

- [1] H. Harashima, K. Aizawa, T. Saito: "Model-Based Analysis Synthesis Coding of Videotelephone Images". Trans. IEICE Japan, E72. 5, pp. 452-459 (May 1989)
- [2] K. Aizawa, Y. Yamada, H. Harashima, T. Saito: "Model-Based Synthesis Image Coding System". Proc. Globecom'87, pp. 2. 3. 1-2. 3. 5 (Nov. 1987)
- [3] K. Aizawa, H. Harashima, T. Saito: "A Model-Based Analysis Synthesis Image Coding Scheme". Trans. IEICE Japan, J72-B. 3, pp. 200-207 (1989-03)
- [4] C. S. Choi, H. Harashima, T. Takebe: "Description and Synthesis of the Facial Expressions for Analysis/Synthesis Image Coding". IEICE Japan Tech. Rep. IE88-84 (Dec. 1988)
- [5] C. S. Choi, H. Harashima, T. Takebe: "3-D Model-Based Synthesis of Facial Expressions and Shape Deformation". Soc. Instrum. Contr. Eng., Human Interface, N&R, 4. 2, pp. 157-166 (May 1989)
- [6] P. Ekman, W. V. Friesen: "Facial Action Coding System". Consulting Psychologist Press (1977)