

表情의 知的 符号化 方式

崔昌石*, 相澤清晴**,
金澤大学 工学部*

原島博**, 武部 幹*,
東京大学 工学部**

A KNOWLEDGE-BASED IMAGE CODING SYSTEM OF FACIAL EXPRESSIONS

Chang-Seok Choi* Kiyoharu Aizawa** Hiroshi Harashima** Tsuyosi Takebe*

The Faculty of Kanazawa University* The Faculty of Tokyo University **

Abstract

A new image coding system called knowledge-based coding is presented which uses positively the knowledge for coding objects. The 3-dimensional model and the psychological knowledge for a person's face can be used as the knowledge for a coding system, when a person's face is coding object.

A transmitter analyzes input images and a receiver synthesizes output images using them. The construction of 3-dimensional model and the introduction of the psychological knowledge to the new coding system are discussed. The basic facial expressions are synthesized from a neutral face using them. The synthesized images are very natural.

필요가 있을 것이다. 좀더 구체적으로 말하면, TV 전화, TV 회의 등에서는 부호화의 대상이 주로 人物像이고, 얼굴에 대한 구조정보를 지식으로 하여, 송신측에서는 입력화상을 분석·인식하여, 정보를 추출, 그 결과만을 전송하고 수신측에서는 구조정보를 변경·합성하는 새로운 통신 방식을 생각할 수 있다. 이 방식에서는 정보의 내용을 분석하여 그 결과만을 부호화하기 때문에, 초저 비트레이트(Ultra bit rate/pixel) 부호화는 물론, 사용자에 적합한 매디어의 선택, 인간과 기계의 인터페이스(man-machine interface)에의 응용도 기대되고 있다.^[1] 이 방식은 지식을 적극적으로 이용하는 측면에서 지식부호화(Knowledge-based Coding)라고 칭한다.

본문에서는 지식부호화 방식을 개관하고, 부호화 대상 모델의 구성 및 표점에 대한 지식을 도입한 부호화 방식에 대해서 논하고자 한다.

1. 序論

현재 주류를 이루고 있는 画像에 대한 高性能符号化方式은 画素에 대한 취도치의 과정에

주목하여 画素間相關이나 振幅分布 등의 統計的 지식을 이용하고 있으나 정보원의 내용에 대한 지식은 이용하고 있지 않다. 이 방식은 적용대상이 넓은 반면, 저비트레이트(Low bit rate/pixel)에서는 부호화 감응이 커져 부자연스런 화상을 제공하게 된다. 지식을 이용하는 측면에서 부호화의 원리적인 구성을 圖1과 같이 생각한다.

한편, 부호화 대상이 한정되어 있는 경우에는 정보원의 내용에 대한 특별한 지식을 적극적으로 이용하여, 통신목적 달성에 필요한 정보만을 추출·부호화하고, 복호시에는 전송된 정보와 정보원에 대한 지식을 이용하여 화상을 합성하는 방식이 생각된다. 이 방식에서는 송수신 양방에 부호화 대상에 대한 지식을 공유할

2. 人物像의 知的 符号化 方式

인물상을 대상으로 한 知的 符号化 方式을 圖2에 나타내고, 간단히 설명한다.

1) 얼굴의 3차원 모델과 표점에 대한 지식.

- 인물상은 원래 3차원의 얼굴을 평면에 투영한 상으로, 얼굴의 3차원 모델을 복원하여 지식으로 이용한다.
- 얼굴에 나타난 특징의지나 감정을 분석·가중·합성하기 위해 표점에 관한 지식을 도입한다.

2) 송신측에서의 분석

- 입력 화상에서 인물과 배경을 분리한다.
- 顔面部의 3차원 운동량(회전·평행이동)을 추정한다.
- 표점에 대한 지식을 이용하여, 표점을 분석·인식한다.
- 모델 및 지식의 불충분한 부분을 학습·갱신한다.

3) 수신측에서의 합성.

- 전송된 정보에 따라 모델을 변형하여, 화상을 합성·

표시한다.

- 모델 및 지식을 갱신한다.
- 모델에 포함되어 있지 않는 정보(배경등)의 생성 표시한다.

지적 부호화 방식에서 전송되는 정보는 입력 화상에 대한 과정 정보가 아니라, 송신측에서 분석한 결과이기 때문에 초저 비트레이트 전송이 가능하고, 수신측에서는 모델과 지식을 이용해서 합성하기 때문에 자연스런 화상을 제공할 수 있다.

3. 얼굴의 3차원 모델의 구성^[2]

인물상의 분석·합성에 있어서 대상에 출현한 3차원 모델은 매우 중요하다. 얼굴의 3차원 모델을 구성하는 과정을 ①~③에 나타낸다. 이 과정에서 대상인물의 정면상(④)을 1매 준비한다. 이후의 합성상은 모두 이 정면상으로부터 합성한 것이다.

1) 얼굴표면 와이어 프레임 모델 (Surface wire-frame model; 이후 WFM라고 약칭한다): 무표정의 마네킨의 모뎀표면에 중첩 등간격으로 십자선을 그려, 얼굴 곡면을 충분히 근사할 수 있는 작은 삼각형으로 분할한다. 이 마네킨을 2방향에서 찍은 사진으로부터 각점의 3차원 좌표를 계산·복원하여, 각점을 연결하여 WFM을 얻는다. (④5)

2) WFM의 변환: WFM의 크기를 대상 인물의 정면상의 대체적인 크기에 맞추기 위하여, WFM의 4특정점(④5의 4특정점)의 인물정면상의 대응점과 일치하도록 어떤 변환(Affine Transform)을 한다. 원뿔 및 투영도 원시시킨다. (④6)

3) 정면상의 WFM에 투영: 정면상이 평행투영이라 가정하여, WFM에 투영한다. 투영은 삼각형 단위로 하되, 삼각형의 양변을 축으로한 좌표계를 사용하여, 画素의 위치 및 휘도치를 추적하여, 얼굴의 3차원 모델을 구성한다.

원래 3차원 모델이 구성되면, 정면상은 물론 여러 가지 각도에서 본 인물상의 합성이 가능하다. (④7) 그러나 얼굴의 3차원 모델은 정면상 1매로부터 구성했기 때문에, 측면에는 불충분한 부분이 생길 수 있어, 갱신을 필요로 하는 경우가 있다. 갱신 방법은 삼각형의 법선 벡터와 투영면과의 각도를 비교하여, 삼각형별로 학습·갱신함으로써, 모든 방향에서 보아도 변경한 3차원 모델을 얻을 수 있다.

4. 表情의 記述 및 合成

얼굴의 3차원 모델이 구성되면, 두부의 운동량은 추적·합성할 수 있지만, 표정을 분석·기술·합성하기 위해서는 표정의 기술 단위가 필요하다. 표정의

기술법으로서 심리학에서 연구되어 있는 FACS(Facial Action Coding System)를 도입하여 기술하는 방법과 더 나아가서 인간의 기본표정을 단위로 기술하는 방법이 생각된다. 이들 기술법과 얼굴 3차원 모델을 결부시켜 부호화에 이용이 가능할 것이다.

1) FACS: FACS란 얼굴의 근육의 수축·이완을 44개의 AU(Action Unit)라는 기본동작으로 분할하여, 인간의 모든 표정을 AU의 조합으로 기술하려는 시스템이다. 예를 들면 AU1은 눈썹의 내측을 올린다는데, AU43은 눈을 감는다는 것이다. FACS의 지식을 도입, 3차원 모델을 변형하는 규칙을 정해 놓으면, 표정은 AU의 번호와 강도로서 부호화가 가능할 것이다. 즉, 송신측에서는 입력 화상을 AU의 번호와 강도로 분석하면, 수신측에서는 AU의 내용에 따라 얼굴의 3차원 모델을 변형하여 합성·표시하면 될 것이다.

2) 기본표정: 심리학에 의하면 인간의 감정에는 놀람, 공포, 혐오, 노여움, 기쁨, 슬픔 등 6개의 기본감정이 있고, 기본감정을 얼굴에 표시하는 기본표정이 있다고 한다. 여러 가지의 표정은 이들 기본표정의 조합으로 생각할 수 있다. FACS의 AU와 기본표정의 관계를 정해 놓고, 송신측에서는 입력 화상을 기본표정으로 분석하여, 수신측에서는 합성하는 과정을 생각할 수 있다.

3) 기본표정의 합성^[3]: 기본표정에 대한 얼굴의 특징적인 변화를 문헌[5]에서 조사해, FACS의 AU와의 관계를 정한 후, ④7의 얼굴의 3차원 모델을 이용하여 합성한 결과가 ④8이다. 이 합성상은 대단히 자연스런 상이고, 본 부호화 방식이 초저레이트임에도 불구하고, 얼마만큼 자연스런 상을 제공할 수 있는가의 가능성이 될 수 있으리라고 생각한다.

기본표정을 단위로 한 부호화와 같이, 의미레벨에 까지 부호화가 가능하다면, 통신시스템은 참으로 필요한 정보만을 추출하여, 이용자에게 최적인 통신 매체를 통해 정보를 가공하여 제공하는 지적 인터페이스도 가능할 것이다.

5. 결론

인물상을 대상으로 한 知的符號化方式에 대해서 개관했다. 얼굴 표정의 分析·合成過程에서 지식으로서 이용되는 얼굴의 3차원 모델의 구성 및 표정 기술법에 대해서 논하고, 이들을 이용하여 기본표정의 합성을 시도했다. 합성상은 매우 자연스런 상을 얻었다.

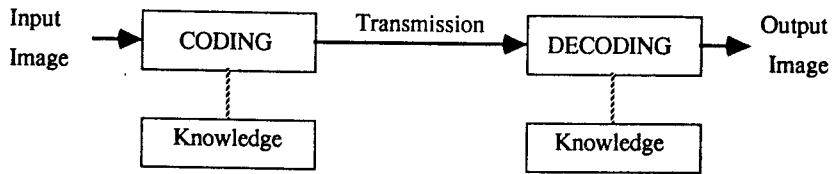


Fig.1 A General Conception of an Image Coding System

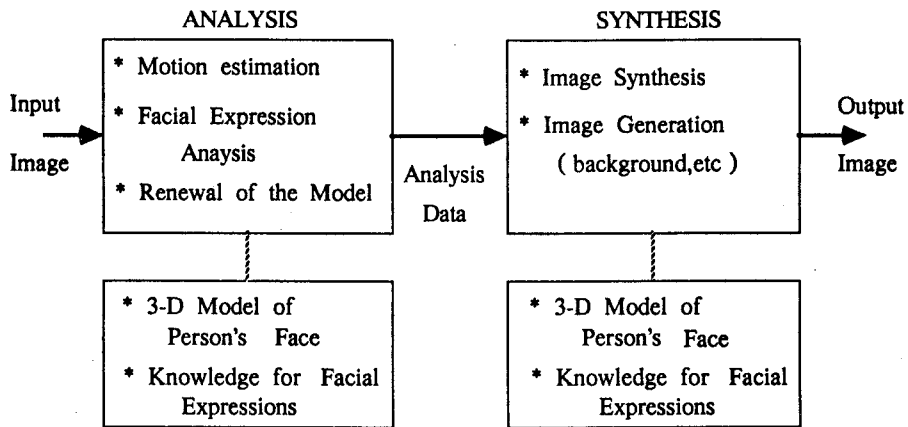


Fig.2 A Knowledge-Based Image Coding System for Facial Expressions

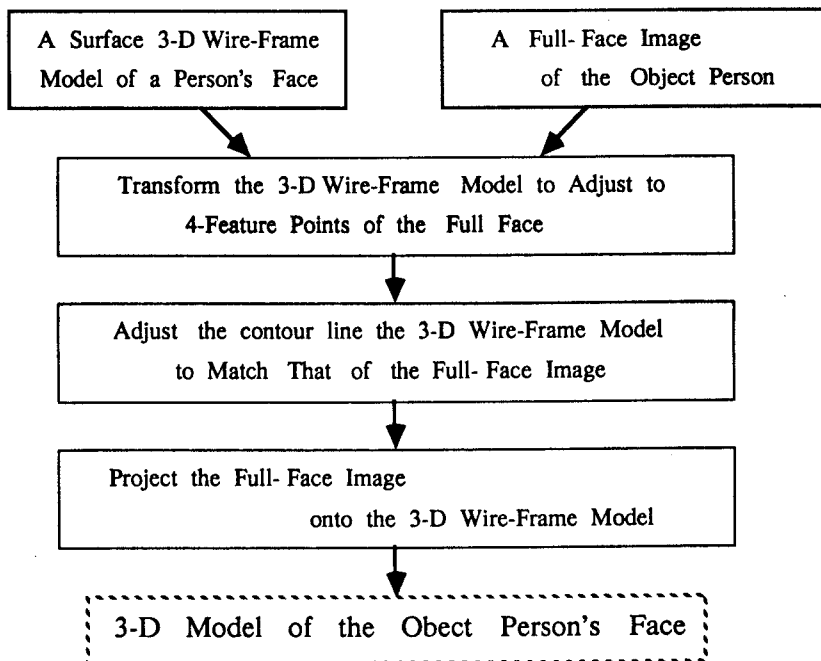


Fig.3 Construction of the 3-D Model of the Person's Face

급후, 부호화 대상의 확장, 특징점의 자동 추출 및 3차원모델의 자동생성, 지식의 체계적인 도입 등에 대해서 검토할 예정이다.

참고문헌

- [1] 原島 外: "次世代 画像符号化의 構想", 信学技報, IE87-1(1987)
- [2] K. Aizawa et al: "Model-based Synthesis Image Coding System", Globecom'87, 45-49, 1987.
- [3] 崔昌石外: "表情的 계층모델과 그構造的記述의 試圖", PCSJ 88, 7-7, 1988
- [4] P. Ekman et al: "Facial Action Coding System", Consulting Psychologist Press (1977)
- [5] 工藤, "表情分析入門", 誠信書房(1987)



Fig.4 A Full Face Image

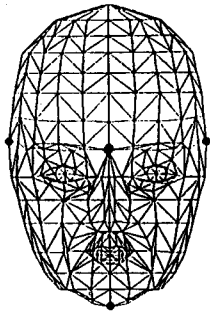


Fig.5 A Surface Wire-Frame Model



Fig.6 The Surface Wire-Frame Model to Adjust a Full Face

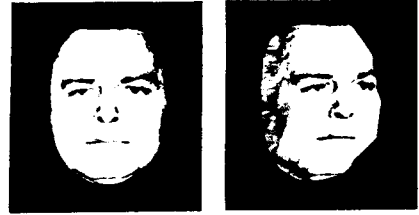


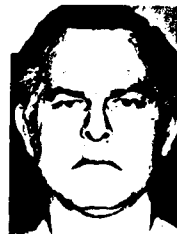
Fig.7 Synthesized Image



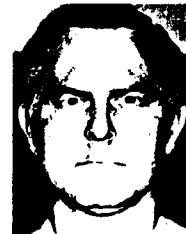
(a) Surprise



(b) Fear



(c) Disgust



(d) Anger



(e) Happiness



(f) Sadness

Fig.8 Synthesized Images of the Basic Facial Expressions