

# 이차원 영상해석을 위한 은닉 마르코프 메쉬 체인 알고리즘

신 봉 기<sup>†</sup>

## 요 약

본 논문에서는 기존의 HMM의 정의와 가정을 확장하여 기존의 단순한 MRF 모델이 아닌 이차원 은닉 마르코프 메쉬 체인형 모델(HMMM)을 제안한다. 체인구조를 위해 상하, 좌우 방향의 부분 순서 관계를 가정하고 또 완전한 격자 조건을 두어 이차원 메쉬의 디코딩 알고리즘, 매개 변수 추정 알고리즘을 제시하였다. 변수 추정 방법은 최우추정법에 의한 EM 방법의 하나로써 반복적으로 최적치에 수렴하는 알고리즘이다. 본 방법은 점진적으로 극소 최적치를 얻을 수 있음이 증명되었다.

## Two-Dimensional Hidden Markov Mesh Chain Algorithms for Image Decoding

Bong-Kee Sin<sup>†</sup>

## ABSTRACT

Distinct from the Markov random field or pseudo 2D HMM models for image analysis, this paper proposes a new model of 2D hidden Markov mesh chain(HMMM) model which subsumes the definitions of and the assumptions underlying the conventional HMM. The proposed model is a new theoretical realization of 2D HMM with the causality of top-down and left-right progression and the complete lattice constraint. These two conditions enable an efficient mesh decoding for model estimation and a recursive maximum likelihood estimation of model parameters. Those algorithms are developed in theoretical perspective and, in particular, the training algorithm, it is proved, attains the optimal set of parameters.

### 1. 서 론

은닉 마르코프 모델(Hidden Markov Model, HMM)은 어떤 사물 패턴의 다양한 변이를 확률적으로 표현하는 통계적 모형이다. HMM은 몇가지 가능한 상태를 기반으로 상태간에 확률적 상태전이로 기술되는 확률 과정인 마르코프 체인, 그리고 마르코프 체인의 확률 함수인 출력과정으로 정의된다. 현재 HMM은 음성인식이나 온라인 필기인식 등 패턴인식 연구 분야에 널리 성공적으로 사용되고 있다. 그런데 HMM은 어떤

상대적 순서가 내제되어 있는 일차원 시계열 신호에 특별히 적합한 모형이며, 본질적인 순서의 개념이 없는 2차원 영상의 경우에는 그대로 적용하기 어렵다. 그러나 최근에는 1D에서의 성공에 자극을 받아 2D 신호에 적합하도록 HMM을 적절히 변형한 2D 마르코프 모델에 대한 연구가 시작되고 있다.

확률 모형으로써 본격적인 2D 영상 모형은 70년대 들어 응용되기 시작한 마르코프 랜덤 필드(Markov Random Field, 이하 MRF) 또는 Gibbs 랜덤 필드(Gibbs Random Field)이다[1-3]. 잠음으로 손상된 영상의 복구[4], 텍스처 모델링[5] 등 하위 수준의 영상 처리 문제에 주로 활용되지만 영상 인식, 언어처리 등 고급 정보

<sup>†</sup> 정 회 원 : 부경대학교 전자컴퓨터정보통신공학부 교수  
논문접수 : 1999년 12월 28일, 심사완료 : 2000년 3월 25일

처리에 이르기까지 다양하게 활용되어 왔다. Gibbs 랜덤 필드[6], 메쉬(mesh) 랜덤 필드[7,8] 등 다양한 변형의 모델과 다양한 방법의 매개변수 추정 방법이 제안되어 있다[9-11]. 그러나 이들 모델은 모두 랜덤 필드에 근거한 것으로써 잡음 제거 등과 같은 국소 영역 작업에는 효과를 발휘하지만, 비선형 공간변형을 모델링 하기에는 기본적인 소수의 매개 변수로는 아무래도 부족하다. 바로 이 점이 MRF가 아직까지 주로 영상 복구, 텍스처 모델 등 하위 수준의 영상 처리 문제에 국한된 이유라고 할 수 있다.

한편 MRF와는 달리 기존의 1D HMM의 선형 전진 구조를 유지한 채 2D로 확장을 시도한 것이 최근에 많이 이용되는 Pseudo-2D HMM(이하 P2D HMM)이다[12, 13]. P2D HMM은 도식상으로 볼 때에만 2차원인 모형이며 실제로는 두단계 계층 구조의 일차원 모형이라고 할 수 있다. 실제로 P2D HMM은 폭 1비트의 수직 프레임 열을 관찰 심볼 또는 벡터로 코딩한 1D HMM과 별로 차이가 없다[14]. P2D HMM은 인체 체 분자 영상과 같이 공간적 변형이 크지 않은 경우에 좋은 효과를 볼 수 있다. 그러나 모션 필드나 필기체 글자처럼 변형이 다양하고 그 정도가 심할 경우에는 효과가 떨어진다[15].

본 논문에서는 HMM의 일차원적 시간 변형, MRF의 2D 국소 변이보다 일반적인 2D 변형을 위한 모형을 제안한다. 우선 2절에서 기존의 HMM을 확장한 2D 은닉 마르코프 메쉬 체인 모델의 개념 및 구성 일반을 정의하고 기존의 MRF 등과의 기본적 차이점을 언급하였다. 그리고 2D 모델의 계산의 효율을 위해서 가정된 순서 체인(causal chain), 완전한 격자 형성 조건을 기반으로 제안 모형의 평가 함수, 마르코프 체인을 구하는 디코딩(decoding)(3절) 목적 함수를 정의하였다. 다음 4절에서는 제안 모형의 매개 변수를 추정하는 EM(estimation-maximization) 알고리즘을 수학적으로 유도하였다. 변수 추정 알고리즘은 반복적으로 이루어지며, 점진적으로 국소 최적치를 얻을 수 있음을 증명하였다. 그리고 5절에서 본 모델의 의미를 다른 각도에서 검토 분석하고 결론에 들어간다

## 2. 2D 은닉 마르코프 메쉬 체인 모델

### 2.1 2D 마르코프 메쉬 체인

주어진 문제는 임의 크기의 영상이 있고 그에 대한

적절한 모델이 있을 때 그 모델의 출력으로 그 영상을 관찰할 확률을 계산하며, 또 궁극적으로는 그 확률이 최대가 되는 모델을 만드는 방법을 찾는 것이다. 우선

$L = \{(u, v) \mid 1 \leq u \leq U, 1 \leq v \leq V\}$ 을  $W = UV$ 개의 사이트(site) 또는 노드를 갖는 이차원 직사각형 격자이고,  $\mathbf{Y}$ 를  $W$ 개의 상태를 2차원 평면 격자 형태로 배열한 메쉬형 마르코프 체인이라고 하자.  $U, V$  각각은 격자의 수직 및 수평 방향의 차원의 크기를 나타낸다. 앞으로 각 상태 또는 노드를 열우선 순으로  $u = 1, \dots, W$ 로 표기하기로 한다.

격자 상에서 각 노드는 주변의 노드에 영향력을 행사할 수 있는데, 이 때 어떤 노드에 영향을 미칠 수 있는 노드의 집합을 이웃이라고 한다. 즉, 노드  $i$ 의 이웃은 일반적으로 다음과 같이 정의한다

**[정의 1]** 다음과 같이 기술되는 격자  $L$ 의 부분 집합의 모임

$$\eta = \{ \eta_i : i \in L, \eta_i \subseteq L \}$$

을 이웃 체계라고 한다. 이는 노드  $i$ 의 이웃  $\eta_i$ 가 다음 두 조건을 만족하는 것과 동치이다.

- 1)  $i \in \eta_i$ ,
- 2)  $j \in \eta_i \rightarrow i \in \eta_j, \quad i = 1, 2, \dots, W$

영상처리 문제에 자주 쓰이는 이웃 체계에는  $\eta^1, \eta^2, \eta^3, \dots$ 의 순으로 계층 구조를 하고 있다.  $\eta^1 = \{ \eta_i^1 \}$ 는 가장 간단한 4-이웃 체계로서, 상하좌우 노드 4개로 구성된 이웃이며, 최근접 이웃(nearest neighbor)이라고도 부른다.  $\eta^2 = \{ \eta_i^2 \}$ 는 가장 가까운 8개의 이웃 노드로 구성된다. 한편 격자의 크기가 유한한 경우 토러스(도넛) 또는 주기적 구조가 아니면 경계 노드의 이웃은 작아진다.

일반적으로 그래프 이론상의 평면 그래프는 그래프 자체의 위상(topology)적 특징을 말하는 것이다. 이와 달리 본 논문에서 제안하려는 모델은 모델의 위상보다는 해석 대상인 영상의 차원과 모델링의 관점에서 본 차원을 의미하고 있다. 따라서 이차원 격자 구조의 마르코프 모델의 명칭으로는 이차원 마르코프 모델보다는 마르코프 메쉬 모델(Markov Mesh Model, MMM) 또는 마르코프 메쉬 체인이 가장 적당하다고 생각된다.

그리고 대부분의 모델에서 실제 계산 알고리즘에 들어가면 순차 계산을 하기 위하여 순서 제약을 가하는 것이 보통이므로 간단히 2D 마르코프 모델 또는 마르코프 체인이라고도 불러도 무방할 것으로 생각된다.

2.2 2-D 은닉 마르코프 메쉬 모델

은닉 마르코프 모델(HMM)의 관찰은 하부의 마르코프 체인의 확률 함수로 정의하는 또 하나의 확률(stochastic) 과정이다. 이 개념을 2-D에 적용해보기로 하자.

2-D 은닉 마르코프 메쉬 모델(HMMM)  $A$ 는 2-D 마르코프 메쉬 모델(MMM)과 MMM의 메쉬 출력 함수로 구성된다. 2-D MMM은  $L$ 개 상태가 부분적으로 중복되어 이루어진 총  $W$ 개 상태의 메쉬형 체인  $Y$ 을 구성한다. 2-D 마르코프 메쉬 모델의 관찰은 직사각형의 영상이라고 가정하자. 각 단위 관찰 벡터, 즉 픽셀은 상하좌우의 픽셀과 인접해 있으며 전체적으로는 하나의 메쉬  $X$ 를 형성한다.  $X = \{x_{uv} \in \Omega : 1 \leq u \leq U, 1 \leq v \leq V\}$ 는  $W = UV$ 개의 픽셀의 2차원 직사각형 영상이다.  $\Omega = \{1, 2, \dots, K\}$ 는 픽셀 값, 즉 명암 또는 색의 값이다. 격자의 경우와 마찬가지로 각 픽셀도  $u = 1, \dots, W$ 와 같이 차례로 번호를 매기기로 한다. 2-D HMMM에서 영상 관찰 과정  $X$ 는 하부의 2-D 마르코프 메쉬 체인  $Y$ 의 함수로 정의한다. 임의 메쉬 체인  $Y$ 가 주어졌을 때  $Y$ 에서 영상  $X$ 를 출력할 확률  $P(X|Y, A)$ 는 각 픽셀을 동시에 관찰할 결합 확률이며  $Y$ 를 조건부로 서로 독립이라고 가정하면

$$\begin{aligned} P(x_1, x_2, \dots, x_W | Y, A) &= P(x_1 | Y, A) P(x_2 | Y, A) \dots P(x_W | Y, A) \\ &= P(x_1 | y_1, A) P(x_2 | y_2, A) \dots P(x_W | y_W, A) \\ &= \prod_{i=1}^W P(x_i | y_i, A) \end{aligned}$$

위 관계 중에서 마지막 식은 1D HMM의 경우와 마찬가지로 매  $u$  지점 상태의 출력 확률의 곱으로 계산된다. 각 상태의 기호 확률 분포는 주어졌다고 가정하면, 그 경우 각  $u$  지점의 상태  $y_u$ , 즉 마르코프 체인  $Y$ 만 결정해주면 위 계산은 끝난다.

한편 메쉬 체인의 동작 구조에 관한 문제가 하나 남아 있다. 일반적으로 순서가 없는 2D 격자의 경우 어

떤 노드의 이웃은  $\eta^1$  또는  $\eta^2$ 로 정의되며 모든 이웃 노드의 상태에 따라 행동이 결정된다. 이와 같은 관계는 역으로 현재 노드가 이웃 노드의 상태에도 같은 영향을 미치게 되고, 모델 전체적으로 특징적인 행동을 보이게 된다. 달리 말하면, 2D 마르코프 모델의 각 노드는 이웃 노드를 조건부로 결정되며 이들 노드들의 메쉬 체인은 일반적 형태의(inhomogeneous) 마르코프 랜덤 필드를 구성한다. 2D HMMM에서 관찰되는 영상  $X$ 는 각 노드의 독립적인 관찰의 집합이다. 2D HMMM이 공간적 변형을 표현하는 능력은 노드 사이를 연결하는 전이의 형태에 따라 결정된다. 그리고 그것은 모델의 구조적 특징을 가름하는 중요한 기준이 된다.

2.3 2-D HMMM의 메쉬 체인

마르코프 성질은 주어진 사건을 결정하는 선행 조건이 존재를 필요로 한다. 그러나 2-D 메쉬 구조에서는 특별한 순서가 없기에 문제점이 발생한다. 위, 아래도 없다. 오직 주위, 즉 이웃만 있을 뿐이다. 이웃의 정의와 노드간의 관계(relation)에 따라서 다음과 같은 세 가지의 극단적인 모델을 들 수 있다. 메쉬 체인을 구성하는 하부 MMM의 상태수를  $L$ 이라고 하자.

첫째, Ergodic 모델.  $p(i|j), i, j = 1, \dots, L$ . 노드  $i$ 의 이웃은  $\eta_i \subseteq L$ 이다. 하부 격자 프로세스 구성에 제약이 없는 가장 일반적인 마르코프 모델로써 ergodic 모델이라고도 한다. 따라서 모델의 구조로 볼 때에는 1D 또는 2D와 같은 차원의 구별은 무의미하다. 동적 프로그래밍 알고리즘을 가질할 때 계산량은  $L^2 W \sim N^6$ 로써 (비슷한 크기의 정방형 격자 및 영상을 가정했을 때) 복잡도가 너무 높아 그대로는 별로 실용성이 없다.

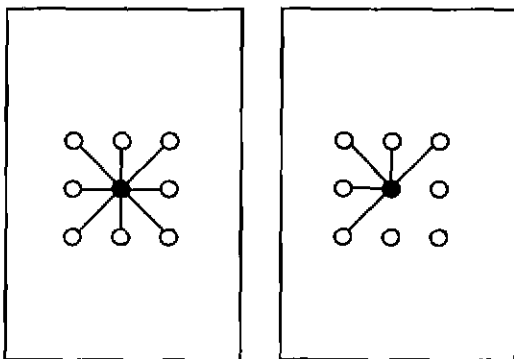
둘째, Noncausal 모델.  $p(i|j) = p(j|i), i, j = 1, \dots, L$  이고  $i \neq j$  부분적으로 대칭적인, 따라서 방향성이 있는 상태 전이를 갖는 noncausal 모델. Ergodic 모델과 마찬가지로 모든  $i = 1, \dots, L$ 에 대하여  $\sum_j p(j|i) = 1$  조건을 만족해야 한다. 그러나 일반적으로 이 확률 조건이 만족되는지 여부와 격자 전체로는 어떤 성질(behavior)을 보이는가 하는 점은 밝혀져 있지 않다. 이 점에 대해서는 아직 직접적으로 연구된 바가 없으며, 수학적 관점에서 흥미 있는 주제라고 판단된다. 이 모델은  $p(j|i) \neq p(j+d|i+d)$ 인 성질 때문에 아래 랜덤 필드 모델과 구분된다.

마지막으로 랜덤 필드 모델  $p(i|\pm d) = p(i\pm d|i)$ ,  $i = 1, \dots, L$ ,  $d$ 는 (양의)정수이고  $i\pm d \in \eta_i$ 를 만족한다. 이 모델은 전체적으로 균일한 특성을 지니는 필드 모델로서 MRF가 여기에 포함된다. 이 경우 노드간 전이는 필요 없고 단지 노드간 공간적 관계에 따른 상관관계, 예컨대, click potential만 있을 뿐이다. 따라서 관찰 영상 데이터만으로 전체의 확률이 결정된다. 이 모형은 엄밀히 말해 HMM은 아니고 MRF에 해당된다. 모형의 확률 분포는 Gibbs 분포를 따른다 [10].

본 논문에서는 첫 번째 부류의 하나로써 다음 소절에서 설명할 가정과 제약 조건을 추가한 모형을 다룬다 여기서 도입하는 가정과 조건은 일차원적 마르코프 체인을 2-D 메쉬형 체인으로 확장하고 2-D 영상에 대한 사람의 직관 및 현실적인 효용성에 근거한 것이다.

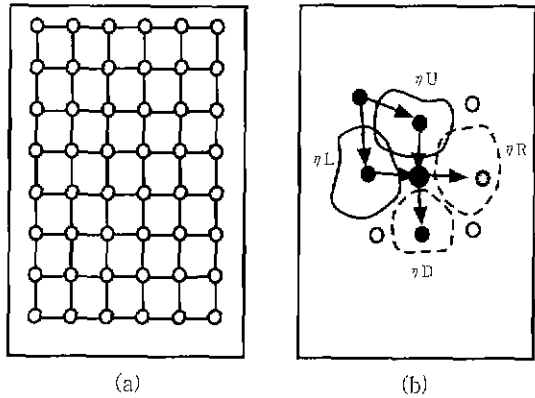
2.4 2D HMMM의 격자 과정

(그림 1)은 마르코프 메쉬의 격자 구조를 나타낸다. (그림 1(a))는 모든 이웃 노드로부터 연결이 되어 있는 구조인데 결과적으로 모델의 구조상으로는 특별한 방향성이 없다. 실제로 시간을 배제한 순수한 공간에는 자연스러운 순서가 없기 때문에 causal model을 사용한다면 그것을 정당화하기는 어렵다. 그러나 계산량을 줄이기 위해서 다음과 같은 두 가지 가정을 도입하였다. 즉, 격자를 구성하는데 있어서 (그림 1(b))와 같이 (1) 노드간 전이를 제한하여 2차(order) 이웃 체계를 가정하였고 (2) 2-D 공간의 경우 위에서 아래로, 왼쪽에서 오른쪽 방향의 순서관계를 가정하였다.



(a) 무방향성의 2D 메쉬      (b) 방향 causal 메쉬

(그림 1) 격자의 연결 구조(위상):



(그림 2) 메쉬 격자 (a) 완전 격자, (b) 격자의 형성과 네 방향의 이웃:  $\eta^U, \eta^L, \eta^R, \eta^D \subseteq L$ .

2-D HMMM의 노드 연결 구조는 일반적인 ergodic 모델과 같이 복잡도가 높은 구조(topology)보다 영상의 2-D 평면성 모델의 구조 자체로 표현하는 것이 바람직하다 그리고 2차원 영상의 다양한 부분적 공간 변형을 쉽게 모델링할 수 있는 8-이웃 구조를 선택하되 4-이웃 구조의 완전한 격자를 구성하도록 한다. 즉 (그림 2)와 같은 구조의 4-이웃 격자를 만든다. 그림의 격자는 2-D 모델로서 보나 자연스럽게 직관적으로 더 적 적절하다.

그림의 구조는, MRF에서 대상으로 하는 단순한 잡음 과정이 아닌 동적 공간 변형(Dynamic Space Warping) 과정을 모델링 한다. 이를 위해서 2-D에서 causality는

$$a_{i|j} = p(y_u = j | y_{u-v} = i), \quad h \in \eta_j^U$$

$$a_{j|}^{\leftarrow} = p(y_u = j | y_{u-1} = i), \quad i \in \eta_j^L$$

의 두 가지로 나누어 표현한다. 여기서  $a_{i|j}^{\leftarrow}$ 는 노드  $h$ 에서 노드  $j$ 로의 수직방향 조건-결과 관계를,  $a_{j|}^{\leftarrow}$ 는 노드  $i$ 에서 노드  $j$ 로의 수평방향 관계를 나타낸다 노드  $j$ 의 위 행에서 노드  $j$ 로 전이 가능한 입력 노드의 집합을  $\eta_j^U$ , 왼쪽에서 전이해 들어올 수 있는 노드의 집합을  $\eta_j^L$ , 그리고 노드  $j$ 에서 전이해 갈 수 있는 오른쪽 및 아래쪽 노드의 집합을 각각  $\eta_j^R, \eta_j^D$  라고 하자. 그러면 다음과 같은 확률 조건을 만족한다.

$$\sum_{k \in \eta_j^u} a_{jk}^{\rightarrow} = 1, \quad a_{ij}^{\rightarrow} \geq 0,$$

$$\sum_{l \in \eta_j^d} a_{jl}^{\downarrow} = 1, \quad a_{ij}^{\downarrow} \geq 0 \quad (1)$$

$\eta_j^R$ 와  $\eta_j^D$ 는 노드  $j$ 에서 본 이웃. 즉 전이가 가능한 이웃이다. 앞으로는 완전 메쉬를 구성하는데 가장 간단하면서도 충분히 일반적인 다음과 같은 이웃 체계를 가정한다.

$$\eta_j^U = \{j-N-1, j-N, j-N+1, j\}$$

$$\eta_j^L = \{j-N-1, j-1, j+N-1, j\}$$

$$\eta_j^R = \{j-N+1, j+1, j+N+1, j\}$$

$$\eta_j^D = \{j+N-1, j+N, j+N+1, j\}$$

그러면 이제 본 논문에서 제안하는 2-D HMM를 정의할 수 있게 되었다.

**[정의 2]** 마르코프 메쉬 체인  $\mathbf{Y}_{1,t}$ 의 확률 함수  $A$ 는 다음과 같은 조건을 만족할 때 2-D 은닉 마르코프 메쉬 체인 또는 2D HMM이다.  $\mathbf{X}_{1,t} = x_1 x_2 \dots x_t$  이고

$\mathbf{Y}_{1,t} = y_1 y_2 \dots y_t$  일 때 :

- (a)  $P(\mathbf{X}_{1,u}, \mathbf{Y}_{1,u} | A) = P(\mathbf{X}_{1,u} | \mathbf{Y}_{1,u}, A) P(\mathbf{Y}_{1,u} | A)$   
(중확률 과정)
- (b)  $P(y_u = j | y_{u-1} = i) \geq 0, \quad i \in \eta_j^L$  (왼쪽 이웃)
- (c)  $P(y_u = j | y_{u-v} = k) \geq 0, \quad k \in \eta_j^U$  (윗쪽 이웃)
- (d)  $\sum_{k \in \eta_j^u} P(k | j) = 1$  (수평 방향 전이),
- (e)  $\sum_{l \in \eta_j^d} P(l | j) = 1$  (수직 방향 전이)

위 정의에 따르면 (b)와 (c)는 각각  $a_{ij}^{\rightarrow}$ 와  $a_{ij}^{\downarrow}$ 를 나타낸다.

2-D 격자에서 위에서 아래로 그리고 좌에서 우로의 진행 관계에 따라 마르코프 성질과 동적 프로그래밍 원리를 적용하면 다음과 같은 전진확률(forward probability)과 후진확률(backward probability)을 얻을 수 있다.

$$\alpha_u(j) = \max_{k \in \eta_j^u, i \in \eta_j^d} a_{ij}^{\rightarrow} a_{jk}^{\downarrow} b_j(x_u) \alpha_{u-1}(i),$$

$$j = 1, \dots, L, \quad u = 1, \dots, W \quad (2)$$

$$\beta_u(j) = \text{MAX}_{k \in \eta_j^d, l \in \eta_j^u} a_{jk}^{\downarrow} a_{jl}^{\uparrow} b_j(x_{u+1}) \beta_{u+1}(k),$$

$$j = L, \dots, 1, \quad u = W, \dots, 1 \quad (3)$$

전진 확률  $\alpha_u(j)$ 는 부분 메쉬 체인  $\mathbf{Y}_{1,j} = y_1 y_2 \dots y_j$ 에서 부분 영상  $\mathbf{X}_{1,u} = x_1 x_2 \dots x_u$ 를 관찰할 확률을 말한다. 후진 확률  $\beta_u(j)$ 는 현재 노드  $j$ , 현재 관찰 데이터  $x_u$  이후 앞으로의 상태 전계를 계산하는 확률로서 부분 메쉬 체인  $\mathbf{Y}_{j+1,L} = y_{j+1} y_{j+2} \dots y_L$ 에서 나머지 관찰 데이터  $\mathbf{X}_{u+1,W} = x_{u+1} x_{u+2} \dots x_W$ 를 출력할 확률을 말한다. 초기 경계 조건은 아래와 같다.

$$\alpha_1(1) = b_1(x_1)$$

$$\alpha_u(j) = \max_{i \in \eta_j^u} a_{ij}^{\rightarrow} b_j(x_u) \alpha_{u-1}(i),$$

$$j = 1, \dots, N, \quad u = 2, \dots, V$$

$$\beta_W(L) = 1 \quad (4)$$

$$\beta_u(j) = \max_{k \in \eta_j^d} a_{jk}^{\downarrow} b_k(x_{u+1}) \beta_{u+1}(k),$$

$$j = L, \dots, L-N+1$$

$$u = W, \dots, W-V+1$$

한편 부분 메쉬 정보, 즉 노드간 체인 구성 정보는

$$(\hat{k}, \hat{l})_u(j) = \arg \max_{k \in \eta_j^u, l \in \eta_j^d} a_{ij}^{\rightarrow} a_{lk}^{\downarrow} b_j(x_u) \alpha_{u-1}(i),$$

$$j = 1, \dots, L, \quad u = 1, \dots, W \quad (5)$$

$$(\hat{k}, \hat{l})_u(j) = \arg \max_{k \in \eta_j^d, l \in \eta_j^u} a_{jk}^{\downarrow} a_{jl}^{\uparrow} b_j(x_{u+1}) \beta_{u+1}(k),$$

$$j = L, \dots, 1, \quad u = W, \dots, 1 \quad (6)$$

로 정의된다. 여기서

$$\text{Left}(j) = \hat{l},$$

$$\text{Up}(j) = \hat{k},$$

$$\text{Right}(j) = \hat{k},$$

$$\text{Down}(j) = \hat{l}$$

이라고 하자. 그러면 위 식 (2), (3)의 계산결과가 완전한 격자가 되기 위해서는

$$\text{Left}(\hat{k}) = \text{Up}(\hat{l}),$$

$$\text{Down}(\hat{k}) = \text{Right}(\hat{l}) \quad (7)$$

와

$$\alpha_u(j) = \max_{k \in \eta_j^u} a_{ij}^{\rightarrow} b_j(x_u) \alpha_{u-1}(R_k),$$

$$j = 1, N+1, 2N+1, \dots, (M-1)N+1,$$

$$u = V+1, 2V+1, \dots, (U-1)V+1$$

$$\beta_u(j) = \max_{l \in \mathcal{Q}^*} a_{jl}^{\downarrow} b_{lk}(x_{u+1}) \beta_{u+1}(L_l),$$

$$j = L, L-N, j-2N, \dots, N \quad (8)$$

$$u = W-V, W-2V, \dots, V$$

의 두 조건을 만족해야 한다 여기서  $R_h = Right^{N-1}(h)$ ,  $L_l = Left^{N-1}(l)$ 로써 격자상에서 각각 노드  $h$ 의 최우단 노드,  $l$ 의 최좌단 노드를 나타내며, 자연수  $n$ 에 대하여 아래와 같은 합성함수로 정의된다.

$$x = Right^0(x) = Left^0(x)$$

$$Right^n(x) = Right(Right^{n-1}(x))$$

$$Left^n(x) = Left(Left^{n-1}(x))$$

식 (7), (8)은 격자 형성을 위해 만족해야 하는 격자 형성 조건(lattice constraint)이다. 이제 (2), (3) 식의 전진 확률과 후진 확률을 이용하면

$$P(\mathbf{X}|A) = \max_{\alpha} \alpha_x(j) \left[ \prod_{k=1}^{V-1} a_{jk}^{\downarrow} \right] \beta_x(j) \quad (9)$$

여기서 메쉬 구조의 격자를 만들기 위해서는

$$h_k = y_{u-v+k} = Right(h_{k-1}) \text{ 그리고 } h_0 = h,$$

$$j_k = y_{u+v} = Right(j_{k-1}) \text{ 그리고 } j_0 = j$$

을 만족해야 한다.

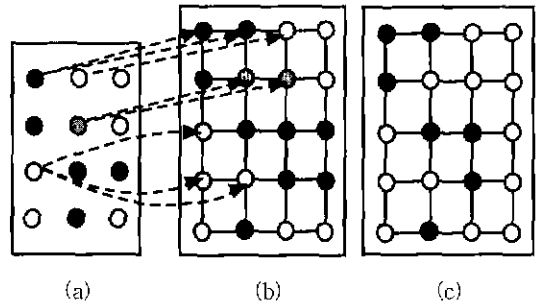
### 3. 마르코프 메쉬 체인 디코딩

HMM과 마찬가지로 2-D HMMM의 관찰 영상  $\mathbf{X}$ 의 출력 과정은 하부 메쉬 체인의 대응 상태에만 종속된, 따라서 출력의 각 픽셀끼리는 서로 독립인 과정으로 정의되어 있다 2-D HMMM의 메쉬 체인을 형성하는 마르코프 모델은  $L$  개의 노드를 가진 격자형 모델로서 부분적으로 중복을 허용하여 크기  $W$ 의 메쉬 영상을 출력한다. 본 논문에서는 암시적으로  $L \leq W$ 로 가정한다. 2D 마르코프 모델의 각 노드의 상태값은 이웃 노드를 조건부로 결정될 뿐만 아니라 위치에도 종속적이므로 마르코프 메쉬 체인은 MRF와 구별된다.

2D HMM  $A$ 의 마르코프 모델을  $\mathcal{Y}$ 이라고 하자.  $A$ 에서 디코딩 문제는 관찰 영상  $\mathbf{X}$ 에 대해 최대 우도의 하부 마르코프 메쉬 체인  $\mathbf{Y}^*$ 를 찾는 것이며

$$f: \mathcal{Y} \rightarrow \mathcal{Y}$$

로 정의된다. 이 함수는 모델  $\mathcal{Y}$ 의 각 노드의 상태를 부분적으로 중복한 배열 방법(warping function)을 나타낸다. (그림 3)을 참조하라.  $\mathbf{Y}^*$ 는 하나의 완전한 격자이다. 그런데 앞 절에서 이미 격자 조건을 적용하여 식 (2)와 같이 전진 확률 함수를 정의하였으므로 그 결과가 곧 디코딩 결과가 된다. 따라서 메쉬 체인은 전진 계산 과정 후 (4)를 이용하여 격자 역추적만 하면 얻을 수 있다.



(그림 3) 디코딩 함수  $f$  (점선 화살표).

(a) 12 노드 마르코프 모델, (b)  $\mathbf{Y}$ , (c) 입력  $\mathbf{X}$ .

### 4. 매개 변수 추정

2D HMMM의 확률 추정식  $P(\mathbf{X}|A)$ 은 대단히 복잡하기 때문에 주어진 샘플의 집합으로부터 최적의 매개변수를 결정할 수 있는 해석적인 방법은 없다. 그러나 주어진 초기 상태에서 더 나은 매개변수의 집합을 구할 수 있는 방법은 가능하다. 2D HMM  $A$ 에서 임의의 마르코프 메쉬 체인을  $\mathbf{Y}$ 라고 하자. 임의의 입력 영상  $\mathbf{X}$ 에 대한 모델의 확률은

$$P(\mathbf{X}|A) = \sum_{\mathbf{Y}} P(\mathbf{X}, \mathbf{Y}|A) \quad (10)$$

로 쓸 수 있다 위 식 우변의 각 항은 입력 영상과 임의의 격자 사이의 결합 확률을 나타낸다. 각 항을 다시 쓰면

$$P(\mathbf{X}, \mathbf{Y}|A) = \prod_{u=1}^W [a_{y_{u-1}y_u}^{\rightarrow} a_{y_{u-1}y_u}^{\downarrow} b_{y_u}(x_u)] \quad (11)$$

와 같이 쓸 수 있다. 위 식에  $\log$ 를 취하면 다음과 같다.

$$\log P(\mathbf{X}, \mathbf{Y}|A) = \sum_{n=1}^W (\log a_{y_{n-1}y_n} + \log a_{y_{n-1}y_n}^{\downarrow} + \log b_{y_n}(x_n)) \quad (12)$$

그러면 2D HMM  $A$ 에 대한  $Q$ -함수는 아래와 같이 전개할 수 있다[16].

$$\begin{aligned} Q(A, \bar{A}) &= \frac{1}{P(\mathbf{X}|A)} \sum_{\mathbf{Y}} P(\mathbf{X}, \mathbf{Y}|A) \log P(\mathbf{X}, \mathbf{Y}|\bar{A}) \\ &= \frac{1}{P} \sum_{\mathbf{Y}=\{y_1, y_2, \dots, y_W\}} P(\mathbf{X}, \mathbf{Y}|A) \times \\ &\quad \sum_{n=1}^W (\log \bar{a}_{y_{n-1}y_n} + \log \bar{a}_{y_{n-1}y_n}^{\downarrow} + \log \bar{b}_{y_n}(x_n)) \\ &= \frac{1}{P} \sum_i \sum_j \sum_n P(\mathbf{X}, y_{n-1}=i, y_n=j|A) \log \bar{a}_{ij} \\ &\quad + \frac{1}{P} \sum_h \sum_j \sum_n P(\mathbf{X}, y_{n-1}=h, y_n=j|A) \log \bar{a}_{hj}^{\downarrow} \\ &\quad + \frac{1}{P} \sum_j \sum_k \sum_n P(\mathbf{X}, y_n=j|A) \log \bar{b}_j(x_n) \quad (13) \end{aligned}$$

$P=P(\mathbf{X}|A)$  이다.  $Q$ -함수는  $\bar{A}$  매개 변수집합 전체를 변수로 하는 함수이다. 식 (13)의 마지막 등호의 우변을 다시 정리하면

$$\begin{aligned} Q(A, \bar{A}) &= \sum_h \sum_j c_{hj} \log \bar{a}_{hj} + \sum_i \sum_j d_{ij} \log \bar{a}_{ij}^{\downarrow} \\ &\quad + \sum_j \sum_k e_{jk} \log \bar{b}_j(x_k) \quad (14) \end{aligned}$$

로 쓸 수 있다. 여기서

$$\begin{aligned} c_{hj} &= \frac{\sum_n P(\mathbf{X}, y_{n-1}=h, y_n=j|A)}{P(\mathbf{X}|A)} \\ d_{ij} &= \frac{\sum_n P(\mathbf{X}, y_{n-1}=i, y_n=j|A)}{P(\mathbf{X}|A)} \\ e_{jk} &= \frac{\sum_{x_n=k} P(\mathbf{X}, y_n=j|A)}{\sum_n P(\mathbf{X}, y_n=j|A)} \end{aligned}$$

이다. 결과적으로 매개변수의 추정식은 다음과 같다

$$\bar{a}_{ij} = \frac{\sum_n P(\mathbf{X}, y_{n-1}=i, y_n=j|A)}{\sum_j \sum_n P(\mathbf{X}, y_{n-1}=i, y_n=j|A)} \quad (15)$$

$$\bar{a}_{ij}^{\downarrow} = \frac{\sum_n P(\mathbf{X}, y_{n-1}=h, y_n=j|A)}{\sum_j \sum_n P(\mathbf{X}, y_{n-1}=h, y_n=j|A)} \quad (16)$$

$$\begin{aligned} \bar{b}_{jk} &= \frac{\sum_{x_n=k} P(\mathbf{X}, y_n=j|A)}{\sum_n \sum_{x_n=k} P(\mathbf{X}, y_n=j|A)} \\ &= \frac{\sum_{x_n=k} P(\mathbf{X}, y_n=j|A)}{\sum_n P(\mathbf{X}, y_n=j|A)} \quad (17) \end{aligned}$$

EM 알고리즘의  $Q$ -함수는  $\bar{A}$ 의 함수로 간주하였다. 위 함수는 1D HMM에 비해 매개 변수가 늘어났지만 기본적으로 1D의  $Q$ -함수와 다르지 않다. 따라서 위 알고리즘은 수렴한다고 말할 수 있다

[정리 1]  $Q(A, \bar{A}) \geq Q(A, A)$ 이면  $P(\mathbf{X}|\bar{A}) \geq P(\mathbf{X}|A)$ 이다. 등식은  $P(\mathbf{X}|\bar{A}) = P(\mathbf{X}|A)$ 일 때 성립한다. 위 정리는 간단히 증명할 수 있다.  $\log$  함수의 concavity에 따라 다음과 같이 쓸 수 있다.

$$\begin{aligned} \log \frac{P(\mathbf{X}|\bar{A})}{P(\mathbf{X}|A)} &= \log \left[ \sum_{\mathbf{Y}} \frac{P(\mathbf{X}, \mathbf{Y}|\bar{A})}{P(\mathbf{X}|A)} \right] \\ &= \log \left[ \sum_{\mathbf{Y}} \frac{P(\mathbf{X}, \mathbf{Y}|A)}{P(\mathbf{X}|A)} \frac{P(\mathbf{X}, \mathbf{Y}|\bar{A})}{P(\mathbf{X}, \mathbf{Y}|A)} \right] \\ &\geq \sum_{\mathbf{Y}} \frac{P(\mathbf{X}, \mathbf{Y}|A)}{P(\mathbf{X}|A)} \log \frac{P(\mathbf{X}, \mathbf{Y}|\bar{A})}{P(\mathbf{X}, \mathbf{Y}|A)} \\ &= Q(A, \bar{A}) - Q(A, A) \quad (18) \end{aligned}$$

$A$ 가  $P(\mathbf{X}|A)$ 의 극점이면  $\bar{A}$ 를 변수로 하는  $Q$ -함수의 극점이며 그 역도 성립한다. 위 식에 따르면 새로이 추정된 모델의 변수  $\bar{A}$ 에 대해 오른쪽 항이 양이면 모델의 출력 확률  $P(\mathbf{X}|A)$ 가 증가하며  $A$ 가 극점이 아닌 경우 반드시 증가한다. ■

위 식 (15)~(17)에 의한 추정식은 반복적으로 이루어지며, 위 정리에 따르면 최적치에 항상 수렴함을 알 수 있다. 본 논문에서 제시한 알고리즘은 정확하다. 그러나 일반적으로 국소 최적치에 수렴하며 주어진 파라미터의 초기값에 크게 좌우된다. 다음 절에는 제안 모델의 특징 및 알고리즘의 특징과 계산량에 대하여 고찰하기로 한다.

### 5. 토의 및 결론

음성신호나 온라인 필기와 같은 일차원 시계열 신호와 일반적인 2차원 영상은 신호에 내재된 순서의 유무

에 따라 차이가 난다. 이차원 영상의 모델로서 제안된 MRF는 전역적 균질성(homogeneity) 조건하에서 상태 간의 관계를 clique potential이란 매개변수로 픽셀 사이에 서로 영향을 미쳐 원 영상을 디코딩 해내려는 모형이다 이와 달리 본 논문에서 제안한 2D HMMM은 필드 모델이 아닌 메쉬형 마르코프 체인을 기반으로 전개되었다. 바로 이 점 때문에 보나 엄밀한 의미에서 1-D HMM을 2-D로 확장한 모델이라고 할 수 있다. 그리고 무엇보다도 2D HMMM은 이차원 영상의 비선형 국소 변형을 간단히 모델링할 수 있는 구조를 갖고 있다.

2D HMMM을 정의하기 위하여 설정한 체인 관계(causality)는 이미 2D 영상 모델링 연구에서 종종 언급된 조건이다. 그리고 본 연구에서 처음 제시하였지만 격자 형성 조건은 직관에 따르는 당연한 귀결의 하나이다. 따라서 본 연구의 모형은 보다 일반적이고 직관에 충실한 모형이라는데 큰 의의가 있다고 본다.

한편 2D HMMM의 구조에서 상태 전이의 범위를 2차 이웃( $n^2$ )으로 제한함으로써 계산량을 대폭 줄일 수 있었다. 관찰 영상의 크기가  $W=U \times V$ 이고 모델의 크기가  $L=M \times N$ 라고 하자. Ergodic 모델과 같은 일반적인 모델 구조를 사용한다면  $O(L^2W) = O(M^2, N^2UV)$ 에컨대 모델과 영상이 같은 크기의 정사각형( $N^2$ )이라면 대략  $O(N^6)$ 에 달하는 비현실적인 계산량이 필요하다. 그러나 제안된 모델 구조로 설계할 경우 복잡도를 훨씬 낮은  $O(LW) = O(MNUV)$ 의 수준으로 낮출 수 있다. 물론 이 정도로도 적은 양은 아니지만 모델의 크기(dimension)가 작다면 여러 가지 영상 해석 문제에 충분히 활용할 수 있는 수준이다. 실제로  $L \ll W$ 인 경우가 보통이므로 계산량은 주로 입력 영상의 크기에 비례하고, 따라서 본 연구의 메쉬형 모형은 활용가능성이 충분하다고 본다.

2D HMMM은 기존의 메쉬형 랜덤 필드 모형들과는 기본적인 가정에서부터 다르다. 이전의 필드형 모델은 다양한 국소 변형과 변이를 모델링 하기에는 너무 단순하다. 그러나 본 논문에 제안된 HMMM은 충분한 양의 샘플에서 반복 추정 과정을 거쳐야 하지만, 늘어난 수만권의 독립적인 매개변수들로 보다 강력한 모델링 능력과 비선형 해석 결과를 얻을 수 있다. 2-D HMMM은 기존의 모델보다 영상 경보에 대해 심도 있는 분석이 요구되는 작업, 예컨대 오프라인 필기체의 분석 및 인식, 비선형 모션 필드의 해석, 영상의 이해 등에 유용하게 활용할 수 있을 것이다.

## 참 고 문 헌

- [1] D. Pickard, "A Curious Binary Lattice Process," J Applied Probability, Vol.14, pp.717-731, 1977.
- [2] J. Woods, "Two-Dimensional Discrete Markovian Fields," IEEE Trans. Info. Theory, Vol.IT-18, No. 2, March 1972.
- [3] A. Jain, "Advances in Mathematical Models for Image processing," Proc. IEEE, Vol.69, pp.502-528, 1981
- [4] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," IEEE Trans. PAMI, Vol.PAMI-6, No.6, Nov. 1984.
- [5] C. Acuna, "Texture modeling using Gibbs distributions," CVGIP : Graphical Models and Image Processing, Vol.54, No.3, pp.210-222, May 1992.
- [6] F. Spitzer, "Markov random field and Gibbs ensembles." Amer. Math. Mon., Vol.78, pp.142-154, 1971.
- [7] L. Kanal, "Markov Mesh Models," Computer Graphics and Image Processing, Vol.12, pp.371-375, 1980.
- [8] H.-S. Park and S.-W. Lee, "A truly 2-D hidden Markov model for off-line handwritten character recognition," Pattern Recognition, Vol.31, No.12, pp.1849-1864, 1998.
- [9] S. Lakshmanan and H. Derin, "Simultaneous Parameter Estimation and Segmentation of Gibbs Random Fields Using Simulated Annealing." IEEE Trans. PAMI, Vol.11, No.8, Aug. 1980
- [10] Y. Zhao, X. Zhuang, L. Atlas and L. Anderson, "Parameter Estimation and Restoration of Noisy Images Using Gibbs Distributions in Hidden Markov Models." CVGIP : Graphical Models and Image Processing, Vol.54, No.3, pp.187-197, May 1992.
- [11] J. Zhang, "The Mean Field Theory in EM Procedures for Blind Markov Random Field Image Restoration," IEEE Trans. Image Processing, Vol. 2, No.1, pp.27-40, Jan. 1993
- [12] O. Agazzi and S. Kuo, "Pseudo two-dimensional hidden Markov models for document recognition," AT&T Technical Journal, Vol.72, No.5, Sep 1993.



- [13] G. Saon, "Cursive word recognition using a random field based hidden Markov model," Int'l J. Doc. Anal. and Recog., Vol.1, No.4, pp.199-208, 1999.
- [14] W. Cho, S.-W Lee and J. H. Kim, "Modeling and recognition of cursive words with hidden Markov models," Pattern recognition, Vol.28, No.12, 1995.
- [15] B.-K. Sin and J. H. Kim, "Ligature modeling for online cursive script recognition," IEEE Trans PAMI, Vol.19, No.6, June 1997.
- [16] L. Baum and J. Eagon, "An inequality with applications to statistical estimation for probabilistic functions of a Markov process and to a models for ecology," Bull. Ann. Math. Stat., Vol.73, pp.360-363, 1967.



### 신봉기

e-mail : bkshin@mail.pknu.ac.kr

1985년 서울대학교 자원공학과  
졸업(학사)

1987년 한국과학기술원 전산학과  
졸업(석사)

1995년 한국과학기술원 전산학과  
졸업(박사)

1987년~1999년 한국통신 멀티미디어연구소

1999년~현재 부경대학교 전자컴퓨터정보통신공학부  
전임강사

관심분야 : 인공지능, 패턴인식, 지능형 에이전트, 정보  
검색