응용통계연구 (2010)
23(2), 305–315

색조영상에서 랜덤결측화소값 대체를 위한 EM 알고리즘 기반 기법

김승구[1]

[1]상지대학교 컴퓨터데이터정보학과

(2010년 02월 접수, 2010년 3월 채택)

요 약

본 논문에서는 색조영상의 R-, G-, B-성분에서 랜덤결측된 화소값들의 대체를 위한 프리퀀티스틱(frequentictic) 기법을 제공한다. 이 기법은 관측영상을 가우시안 마코프 랜덤필드 상의 실현치로서 가정하고, 주어진 화소 내의 근방화소들이 에지 강도에 따른 서로 다른 분산을 가지는 정규분포를 따른다고 설계함으로써 에지에서 결측화소 대체값이 이질적 색상에 영향 받지 않도록 한다. 이러한 모형하에서 우도가 최대화하도록 결측화소값들을 근사 EM 알고리즘에 기반 한 방법으로 모수들을 추정하고 결측화소를 대체한다. 제안된 방법의 결과들은 보간법에 기초한 대체법과 비교하여 그 유효성을 보인다.

주요용어: 다중 결측, 결측치 대체, 가우시안 마코프 랜덤 필드, EM 알고리즘, 색조영상.

1. 서론

색조영상(RGB color image)의 색상은 R(적), G(녹), B(청) 3 성분의 밀도값으로 결정된다. 현실에서 영상의 일부 화소에서 이 세 성분 중 일부 혹은 전부가 랜덤하게 결측(random missing)되는 경우가 흔하다. 예를 들면, CCD(charged couple device) 카메라는 화소별로 오직 1개의 성분만 기록하고 주변화소들과 결합하여 색상을 출력한다. 그런데 많은 경우 결측화소값 대체(imputation)가 만족스럽지 못한 주된 이유 중 하나는 '에지 화소'들(edge pixels)에서 대체가 정확히 이루어지기가 어려운데 일반적인 영상에는 에지가 너무 많기 때문이다. 에지 화소란 영상 내에 색상이 다른 두 동질적인 면이 만나는 경계영역에 위치하고 있는 화소를 말한다. 이런 화소에 결측된 성분이 있으면 주변화소의 도움으로 대체한 성분값은 이질적 색상 성분에 영향을 받게 된다. 이 경우 명암도 영상과 같이 단변량 영상에서는 시각적으로 그렇게 큰 위화감이 들지는 않는다. 그러나 색조영상의 경우 (적, 녹, 청) 세 성분 중 한 성분값을 그릇되게 추정하여 약간의 변화가 생기면 색상의 큰 차이를 보이게 된다. 이러한 면에서 명암도 영상의 결측치 대체 (Kim 등, 2006; Ogawa 등, 2006)보다 색조 영상의 그것이 더 어렵다. 이후 본 논문에서는 에지 화소에서의 이질적 주변값을 효과적으로 배제하면서 대체하는 방법에 집중되어있다

본 연구에서는 주어진 영상을 MRF(Markov random field)의 실현치로 보고 문제에 접근한다. Blanchet와 Vignes (2009)는 마이크로어레이 영상 결측화소 문제에서 그리고 Dass와 Nair (2003)는 CCD 카메라의 화질 개선을 위해 MRF 모형을 가정하고 베이지안(Bayesian) 접근법으로 결측 화소 대체문제를 다루었다. 최근 김승구 (2009)는 HMRF(hidden MRF) 하에서 정규혼합모형(normal mixture

본 연구는 2009년도 상지대학교 교내연구비 지원에 의해 수행되었음.

[1](220-702) 강원도 원주시 우산동 660 상지대학교 컴퓨터데이터정보학과, 교수. E-mail: sgukimg@sangji.ac.kr

model)을 기반으로 하여 EM (expectation-maximization)을 적용하여 순수 프리퀀티스틱(frequentic-tic) 접근법으로 이 문제를 해결하려고 시도하였으나 결측화소의 비율이 50% 이상인 경우 만족스러운 결과를 얻지 못하는 제한점을 가지고 있다. 이에 대한 개선 방안으로 본 연구에서는, 새롭게 정의한 가우시안 MRF 하에서 EM 알고리즘으로 해결하는 방법을 제공한다. 이때 Dempster 등 (1977)의 순수 EM 알고리즘을 다소 변형한 방법이 적용된다. 제안된 방법은 90% 이상의 결측 비율에서도 매우 양호한 결과를 보여줄 것이다.

이를 위해, 다음절에서는 제안된 방법의 기반 모형을 정의하며, 3절에서는 모형 추정과 결측화소 대체를 위한 변형된 근사 EM 알고리즘을 제공하고, 4절에서는 주어진 영상에 결측비율을 증가시키면서 제안된 방법을 스프라인 보간법(spline interpolation)을 이용한 대체 결과와 비교하면서 그 유효성을 보인다. 5절에서는 결론과 기타 제한점을 정리하였다.

2. 제안된 방법

2.1. 모형 정의

영상 $\boldsymbol{Y} = \{\boldsymbol{Y}_i\}$는 공간과정(spatial process)으로서 공동의 근방계 ∂를 가지는 MRF로 가정하며, $\boldsymbol{y}$는 MRF의 하나의 실현치라 가정한다. 본 연구의 실험에서 ∂는 2차 근방계를 사용하였다. 즉, 한 화소에 대해 8개의 근방화소를 가진다(동서남북 방향 근방화소 4개 그리고 대각선 방향 근방화소 4개). 그리고 $\boldsymbol{y}_i$는 주변적으로(marginally) (3×1) 평균(벡터) $\boldsymbol{\mu}_i$ 및 (3×3) 공분산(행렬) $\boldsymbol{\Sigma}_i$를 가지는 (3-변량) 정규분포 $\mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$를 따른다고 가정한다. 또한 화소 i가 주어졌을 때 근방화소 $j(\in \partial_i)$들에 부과한 어떤 부여값 $e_{ij} = (\boldsymbol{e}_i)_j$를 고려한다. 이때 $e_{ij} = e(\boldsymbol{y}_i, \boldsymbol{y}_j)$는 $e_{ij} > 0$, $e_{ij} = e_{ji}$ 및 $e_{ii} = \infty$를 만족하는 측도로서, 화소값 $\boldsymbol{y}_i$와 대응하는 근방화소값 $\boldsymbol{y}_j (j \in \partial_i)$ 사이의 유사성 정도를 측정한다. e_{ij}에 대해서는 다음 소절에서 자세히 정의할 것이다.

확률변수 벡터 $\boldsymbol{e} = (\boldsymbol{e}_1^T, \ldots, \boldsymbol{e}_n^T)^T$는 확률밀도 $p(\boldsymbol{e})$를 가진다고 할 때, $\boldsymbol{e} = \boldsymbol{e}(\boldsymbol{y})$의 확률적 성질은 $\boldsymbol{y}$만 의존하기 때문에 결합밀도 $h(\boldsymbol{e}, \boldsymbol{y}) = f(\boldsymbol{y})$의 성질을 가진다. 또한 본 연구에서는

$$f(\boldsymbol{y}; \boldsymbol{\theta}) = h(\boldsymbol{y}, \boldsymbol{e}; \boldsymbol{\theta}) \equiv p(\boldsymbol{e}) f_{\boldsymbol{\theta}}(\boldsymbol{y}|\boldsymbol{e}) \tag{2.1}$$

와 같이 모수 $\boldsymbol{\theta} = \{\{\boldsymbol{\mu}_i\}, \{\boldsymbol{\Sigma}_i\}\}$에 대한 정보는 오직 $f_{\boldsymbol{\theta}}(\boldsymbol{y}|\boldsymbol{e})$에만 포함되어 있다고 가정한다. 이러한 조건하에서 식 (2.1)에 대응한 모수 $\boldsymbol{\theta}$에 대한 최우추정문제는 $p(\boldsymbol{e})$와 무관하게 될 것이다.

아울러 대부분의 접근법 (Besag, 1975; Qian과 Titterington, 1992; 등)에서 그러하듯이

$$f_{\boldsymbol{\theta}}(\boldsymbol{y}|\boldsymbol{e}) \approx \prod_{i=1}^{n} f_{\boldsymbol{\theta}_i}(\boldsymbol{y}_i|\boldsymbol{y}_{\partial_i}, \boldsymbol{e}_i)$$

와 같이 각 근방계 화소값들이 주어졌을 때는 화소값들은 서로 독립이라 가정한다(조건부 독립성 가정). 그리고 본 논문에서는

$$f_{\boldsymbol{\theta}_i}(\boldsymbol{y}_i|\boldsymbol{y}_{\partial_i}, \boldsymbol{e}_i) = \prod_{j \in \partial_i} \phi\left(\boldsymbol{y}_j; \boldsymbol{\mu}_j, \frac{\boldsymbol{\Sigma}_j}{e_{ij}}\right), \quad i = 1, \ldots, n \tag{2.2}$$

으로 정의한다. 여기서 $\phi(\boldsymbol{y}_j; \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j/e_{ij})$는 평균 $\boldsymbol{\mu}_j$ 및 공분산 $\boldsymbol{\Sigma}_j/e_{ij}$을 가지는 정규분포 밀도이다.

식 (2.2) 설계의 의도를 좀 더 부연 설명하자면 다음과 같다. 우선, 화소 i가 주어진 조건 하에서 대응하는 근방화소값들은 평균은 유지하지만 부여치 e_{ij}에 의해 서로 다르게 '척도화된 공분산'을 가진 정규분포 $\mathcal{N}(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j/e_{ij})$를 (독립적으로) 따르는 것으로 하였다. 이제, 색상이 거의 동질적인 영역의 화소 i가

주어졌을 때(즉, i가 에지화소가 아닐 때) 근방화소의 $\boldsymbol{y}_j(j \in \partial_i)$들은 $\boldsymbol{y}_i$와 거의 비슷하여 모두 e_{ij}가 클 것이다. 이것은 각 정규밀도가 평균에 집중하게 하여 우도에 각 평균 색상 $\boldsymbol{\mu}_j$에 대한 정보를 고르게 제공한다. 반면, 근방화소 중 하나 j^*가 매우 이질적 색상 영역에 놓여 있다면(즉, i가 에지화소 일 때) $\boldsymbol{y}_{j^*}$와 $\boldsymbol{y}_i$의 차이는 커서 e_{ij^*}는 매우 작은 값을 가질 것이다. 이것은 $\mathcal{N}(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j/e_{ij^*})$를 평평한 밀도로 만들어 예외적으로 우도에 기여하지 못하게 한다. 이와 같이 식 (2.2)의 설계를 통해 어떤 화소가 에지에 위치하였을 경우, 화소 i와 동질적인 색상을 가진 근방화소들만이 영향 미치도록 하는 '선택적 우도'를 구성하게 될 것이다.

이와 같은 가정과 정의 하에서 식 (2.2)를 식 (2.1)에 대입하면 제안된 모형은

$$f(\boldsymbol{y};\boldsymbol{\theta}) \approx \prod_{i=1}^{n} \prod_{j \in \partial_i} \phi\left(\boldsymbol{y}_j; \boldsymbol{\mu}_j, \frac{\boldsymbol{\Sigma}_j}{e_{ij}}\right) \tag{2.3}$$

와 같이 매우 단순한 형태를 취하고 있음을 알 수 있다.

2.2. 부여치 함수

본 연구에서는 두 화소값 벡터의 유사성의 크기 측정하는 함수를

$$e_{ij} = e(\boldsymbol{y}_i, \boldsymbol{y}_j) = \nu_j \times \frac{b}{d(\boldsymbol{y}_i, \boldsymbol{y}_j)^a + c}, \quad j \in \partial_i;\ i = 1, \ldots, n \tag{2.4}$$

로서 정의한다. 여기서 ν_j는 근방화소 위치에 따른 가중치로서, 예를 들어 2차 근방계의 경우 동서남북 근방화소에 대해서는 1 그리고 대각선 근방화소에 대해서는 $1/\sqrt{2}$이다. 또한 a, b, c는 함수의 매개변수인데, 특히 a를 크게 하면 두 화소 간의 유사성을 엄격하게 통제한다. 본 연구의 실험에서는 $a = 1/2$, $b = 1$, $c = 0.01$로 고정하여 사용하였다. 그리고 $d(\boldsymbol{y}_i, \boldsymbol{y}_j)$는 두 화소값 벡터의 거리함수로서 다양한 정의가 가능하지만, 본 연구에서는

$$d(\boldsymbol{y}_i, \boldsymbol{y}_j) = (\boldsymbol{y}_i - \boldsymbol{y}_j)^T \boldsymbol{\Gamma}^{-1} (\boldsymbol{y}_i - \boldsymbol{y}_j)$$

와 같이 Mahalnobis 거리 제곱을 추천한다. 왜냐하면, 저자의 경험 상 많은 거리 측도 중에 Mahalnobis 거리가 시각적 색상 차이를 가장 잘 반영한다고 보기 때문이다. 단, 식 (2.2)에서 $\boldsymbol{\Gamma}_i$는 화소들의 공통의 공분산을 나타낸다. 이를 테면, (일종의 평균 차이 t-통계량에 대한 해석과 유사하게) 두 화소의 색상 간 유클리드 거리가 크더라도 잡음 등의 원인으로 성분들의 분산이 클 때 색상의 차이에 대한 해석을 유보시키며, 반대로 두 화소간 색상 차이가 작더라도 분산이 작으면 거리 차이를 크게 해석하는 것이다. 또한 상관 구조에 따라서도 색상 간 차이에 대한 측정값을 다소 완화시키거나 증대시킨다.

이 함수는 원래 Dass와 Nair (2003)가 사용했던 것을 변형한 것인데, 그들은 ν_j 인자가 없고 거리 함수 $d(\boldsymbol{y}_i, \boldsymbol{y}_j)$는 유클리드 거리 제곱을 사용하였다.

2.3. 결측치 화소 표기

주어진 영상은 $\boldsymbol{y} = (\boldsymbol{y}_1^T, \ldots, \boldsymbol{y}_n^T)^T$과 같이 n개의 화소값들로 이루어져 있으며, 각 화소값은 (3×1) 벡터 $\boldsymbol{y}_i = (y_{i1}, y_{i2}, y_{i3})^T = (r_i, g_i, b_i)^T$과 같이 (적, 녹, 청) 세 성분으로 색상을 정의한다. 여기서 각 성분값은 연속형 자료로 취급한다. 이제 각 화소 $\boldsymbol{y}_i = (y_{i1}, y_{i2}, y_{i3})^T$에서 성분 중 일부 혹은 3-성분값 전부가 임의로 결측되거나 혹은 모두 관측된 상황을 고려하자. 이때 i번째 화소값 벡터를 관측된 성분 ('obs'로 표시)과 결측된 성분 ('mis'으로 표시)으로 나누어 $\boldsymbol{y}_i = (\boldsymbol{y}_i^{\mathrm{obs}^T}, \boldsymbol{y}_i^{\mathrm{mis}^T})^T$ 그리고 영상은

$\boldsymbol{y} = (\boldsymbol{y}^{\text{obs}T}, \boldsymbol{y}^{\text{mis}T})^T$와 같이 나타내겠다. 이렇게 나타내어도 본 논문에서 일반성을 잃지 않는다. 만약 화소 i에 결측이 없다면 $\boldsymbol{y}_i^{\text{mis}} = [\,]$일 것이고, 모든 성분이 결측되었다면 $\boldsymbol{y}_i^{\text{obs}} = [\,]$일 것이다. 이때 우리의 목표는 관측된 화소자료 $\boldsymbol{y}^{\text{obs}}$를 바탕으로 결측 화소의 성분값 $\boldsymbol{y}^{\text{mis}}$을 대체하는 것이다. 아울러 앞으로 $\boldsymbol{y}_i^{\text{obs}}$와 $\boldsymbol{y}_i^{\text{mis}}$의 크기에 따른 분할을 평균 벡터 $\boldsymbol{\mu}_i$와 공분산 행렬 $\boldsymbol{\Sigma}_i$의

$$\boldsymbol{\mu}_i = \begin{pmatrix} \boldsymbol{\mu}_i^{\text{obs}} \\ \boldsymbol{\mu}_i^{\text{mis}} \end{pmatrix}, \quad \boldsymbol{\Sigma}_i = \begin{pmatrix} \boldsymbol{\Sigma}_i^{\text{oo}} & \boldsymbol{\Sigma}_i^{\text{om}} \\ \boldsymbol{\Sigma}_i^{\text{mo}} & \boldsymbol{\Sigma}_i^{\text{mm}} \end{pmatrix}$$

과 같이 표시하자.

3. 결측화소값 대체를 위한 EM 알고리즘

3.1. 표준 EM 적용의 문제점

이제 우리는 관측 화소자료 $\boldsymbol{y}^{\text{obs}} = (\boldsymbol{y}_1^{\text{obs}T}, \ldots, \boldsymbol{y}_n^{\text{obs}T})^T$를 불완비자료, 미관측 $\boldsymbol{y}^{\text{mis}} = (\boldsymbol{y}_1^{\text{mis}T}, \ldots, \boldsymbol{y}_n^{\text{mis}T})^T$을 결측자료 그리고 $\boldsymbol{y} = (\boldsymbol{y}^{\text{obs}T}, \boldsymbol{y}^{\text{mis}T})^T$를 완비자료로 취급한다. 이때 식 (2.3)에 대응하는 완비자료에 대한 로그-우도는 (상수항은 제외하고)

$$\log L\left(\boldsymbol{\theta}; \boldsymbol{y}^{\text{obs}}, \boldsymbol{y}^{\text{mis}}\right) = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j\in\partial_i}|\boldsymbol{\Sigma}_j| - \frac{1}{2}\sum_{i=1}^{n}\sum_{j\in\partial_i} e_{ij}(\boldsymbol{y}_i - \boldsymbol{\mu}_j)^T\boldsymbol{\Sigma}_j^{-1}(\boldsymbol{y}_i - \boldsymbol{\mu}_j) \tag{3.1}$$

와 같이 주어진다. 이때 EM 알고리즘은 $(k+1)$번째 단계의 E-step에서

$$Q\left(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)}\right) = E\left[\log L\left(\boldsymbol{\theta}; \boldsymbol{Y}^{\text{obs}}, \boldsymbol{Y}^{\text{mis}}\right)|\boldsymbol{y}^{\text{obs}}, \boldsymbol{\theta}^{(k)}\right]$$

를 계산하고, M-step에서 $\boldsymbol{\theta}^{(k+1)} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)})$를 구하는 반복과정이다.

만약 e_{ij}가 상수라면 표준적인 EM 알고리즘을 적용하여, E-step에서 모수가 $\boldsymbol{\theta}^{(k)}$일 때 식 (3.1)의 우변의 두번째 항을 구성하는 충분통계량 $\boldsymbol{Y}_i$와 $\boldsymbol{Y}_i\boldsymbol{Y}_i^T$의 $\boldsymbol{y}^{\text{obs}}$에 대한 조건부 기대값을 구하면서 쉽게 해결하게 된다. 그러나 주어진 문제에서 e_{ij}는 상수가 아니라 $(\boldsymbol{Y}_i, \boldsymbol{Y}_{j\in\partial_i})$의 함수로서 확률변수이다. 결국 확률변수 $E_{ij}\boldsymbol{Y}_i$와 $E_{ij}\boldsymbol{Y}_i\boldsymbol{Y}_i^T$의 조건부 기대값을 구해야 하는데, 이 두 변량의 $\boldsymbol{y}^{\text{obs}}$에 대한 조건부 분포를 파악한다는 것은 거의 불가능해 보이기 때문에 E-step을 수행할 수가 없다. 그래서 이 문제를 회피하는 대안을 고려해야만 한다. 우선 생각해 볼 수 있는 두 가지 방법은 다음과 같다. 즉,

(1) 부여치를 계산할 때 관측된 자료만을 이용한다. 즉, $e_{ij} = e(\boldsymbol{y}_i^{\text{obs}}, \boldsymbol{y}_j^{\text{obs}})$를 사용하는 것이다. 이때 두 조건부 기댓값은 각각 $E(E_{ij}\boldsymbol{Y}_i|\boldsymbol{y}^{\text{obs}}, \boldsymbol{\theta}^{(k)}) = e_{ij}E(\boldsymbol{Y}_i|\boldsymbol{y}^{\text{obs}}, \boldsymbol{\theta}^{(k)})$ 및 $E(E_{ij}\boldsymbol{Y}_i\boldsymbol{Y}_i^T|\boldsymbol{y}^{\text{obs}}, \boldsymbol{\theta}^{(k)}) = e_{ij}E(\boldsymbol{Y}_i\boldsymbol{Y}_i^T|\boldsymbol{y}^{\text{obs}}, \boldsymbol{\theta}^{(k)})$가 되어 표준적인 EM 알고리즘으로 해결가능할 것이다. 그러나 i번째 화소에서 관측된 성분이 $\boldsymbol{y}_i = (r_i, \cdot, \cdot)^T$과 같이 적색성분이고 j번째 화소는 $\boldsymbol{y}_j = (\cdot, \cdot, b_j)^T$과 같이 청색성분이라 할 때, 서로 특성이 다른 두 색상 성분의 차이로 색상 유사성의 크기 e_{ij}를 결정하는 것은 비현실적이다. 더구나 많은 화소에서 모든 성분이 결측될 수 있으므로 이런 경우에는 e_{ij}를 얻을 수 없게 된다. 따라서 이 방법은 색조영상 결측치 대체 문제에서는 적용하기 어렵다.

(2) Wei와 Tanner (1990)의 몬테카를로 EM(MCEM) 알고리즘을 적용한다. 이 방법은 주어진 모수 추정값 $\boldsymbol{\theta}_i^{(k)} = \{\boldsymbol{\mu}_i^{(k)}, \boldsymbol{\Sigma}_i^{(k)}\}$하에서 모의실험을 통해 두 기댓값의 경험적 평균을 계산하는 방법이다. 그러나 우리 문제에서는 주어진 모수쌍 $\boldsymbol{\theta}_i^{(k)}$의 개수가 화소의 개수와 같으므로 EM 알고리즘의 매 단계마다 n번의 모의실험을 해야 한다. 따라서 이 방법은 매우 많은 모수를 가지는 영상응용문제에는 처리시간상 적용하기가 어렵다.

따라서 위 두 방법은 고려의 대상에서 제외하고, 다음 소절에서 다른 대안을 모색하도록 한다.

3.2. 근사 E-step

식 (3.1)의 완비자료에 대한 로그-우도를 잠시 $\log L(\boldsymbol{\theta}; \boldsymbol{y}^{\text{obs}}, \boldsymbol{y}^{\text{mis}}) \overset{\text{let}}{=} \ell(\boldsymbol{y}, \boldsymbol{e}; \boldsymbol{\theta})$라 쓰기로 하자. 이때

$$Q\left(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)}\right) = E\left[\ell\left(\boldsymbol{Y}, \boldsymbol{E}; \boldsymbol{\theta}\right)|\boldsymbol{y}^{\text{obs}}, \boldsymbol{\theta}^{(k)}\right] = \int_{\boldsymbol{e}}\int_{\boldsymbol{y}} \ell\left(\boldsymbol{y}, \boldsymbol{e}; \boldsymbol{\theta}\right) h_{\boldsymbol{\theta}^{(k)}}\left(\boldsymbol{y}, \boldsymbol{e}|\boldsymbol{y}^{\text{obs}}\right) d\boldsymbol{y}d\boldsymbol{e}$$
$$= \int_{\boldsymbol{e}}\int_{\boldsymbol{y}} \ell\left(\boldsymbol{y}, \boldsymbol{e}; \boldsymbol{\theta}\right) f_{\boldsymbol{\theta}^{(k)}}\left(\boldsymbol{y}|\boldsymbol{y}^{\text{obs}}, \boldsymbol{e}\right) p(\boldsymbol{e})d\boldsymbol{y}d\boldsymbol{e}$$
$$= \int_{\boldsymbol{e}} Q_{\boldsymbol{e}}\left(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)}\right) p(\boldsymbol{e})d\boldsymbol{e} = E\left[Q_{\boldsymbol{e}}\left(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)}\right)\right]$$

임을 알 수 있다. 단,

$$Q_{\boldsymbol{e}}\left(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)}\right) \overset{\text{let}}{=} \int_{\boldsymbol{y}} \ell\left(\boldsymbol{y}, \boldsymbol{e}; \boldsymbol{\theta}\right) f_{\boldsymbol{\theta}^{(k)}}\left(\boldsymbol{y}|\boldsymbol{y}^{\text{obs}}, \boldsymbol{e}\right) d\boldsymbol{y} = E\left[\log L\left(\boldsymbol{\theta}; \boldsymbol{Y}^{\text{obs}}, \boldsymbol{Y}^{\text{mis}}\right)|\boldsymbol{y}^{\text{obs}}, \boldsymbol{e}, \boldsymbol{\theta}^{(k)}\right] \quad (3.2)$$

을 나타내며, $\boldsymbol{e}$의 적분 영역은 $\mathcal{E} = \{\boldsymbol{e} : \boldsymbol{e} = \boldsymbol{e}(\boldsymbol{y})\}$이다. 식 (3.2)로부터 $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)})$는 $\boldsymbol{e}$에 관하여 $Q_{\boldsymbol{e}}(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)})$의 평균임을 알 수 있다. 만약 $p(\boldsymbol{e})$로부터 임의로 추출한 한 점 $\boldsymbol{e} = \boldsymbol{e}^*(\in \mathcal{E})$가 주어졌을 때 $Q_{\boldsymbol{e}^*}(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)})$가 평균 $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)})$와는 차이가 있겠지만 확률밀도 $p(\boldsymbol{e})$를 고려할 때 (아주 평평한 밀도가 아니라면) 터무니 없이 큰 차이가 나지는 않을 것이다. 더욱이 평균과 가장 가깝도록 하는 $\mathcal{E}$ 상의 어떤 점들 중 하나는 아마 $\boldsymbol{e} = \boldsymbol{e}^{(k)}$ 즉 전 단계에서 얻은 예측값일 것이라 기대할 수 있다. 따라서 본 연구에서는 근사적으로 E-step에서 $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)})$ 대신 식 (3.2)에 $\boldsymbol{e} = \boldsymbol{e}^{(k)}$가 주어진

$$Q_{\boldsymbol{e}^{(k)}}\left(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)}\right) = E\left[\log L\left(\boldsymbol{\theta}; \boldsymbol{Y}^{\text{obs}}, \boldsymbol{Y}^{\text{mis}}\right)|\boldsymbol{y}^{\text{obs}}, \boldsymbol{e}^{(k)}, \boldsymbol{\theta}^{(k)}\right] \approx Q\left(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)}\right) \quad (3.3)$$

를 계산하고, M-step에서 식 (3.3)를 $\boldsymbol{\theta}$에 관하여 최대화하는 알고리즘을 적용할 것이다. 우리의 문제에서 $Q_{\boldsymbol{e}^{(k)}}(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)})$는 $\boldsymbol{\theta}$에 관하여 완전 오목(strictly concave)이므로 반복과정은 단조 수렴하게 된다. 그러나 그 수렴점이 원래의 $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)})$를 최대화할 때와 일치하는지 본 논문에서 이론적으로 확인하지는 못하였다. 그럼에도 불구하고 제안된 알고리즘은 4절의 실험을 통해 항상 매우 의미있는 점으로 수렴하는 결과를 보였다.

이제 제안된 방법의 (근사) EM 알고리즘은 다음과 같은 과정으로 수행한다.

주어진 초기 추정치 $\boldsymbol{\theta}_i^{(0)} = \{\boldsymbol{\mu}_i^{(0)}, \boldsymbol{\Sigma}_i^{(0)}\}$ 및 초기 대체값 $\boldsymbol{y}^{(0)} = (\boldsymbol{y}^{\text{obs}^T}, \boldsymbol{y}^{\text{mis}(0)^T})^T$에 대해, E-step에서 먼저 초기 대체값 $\boldsymbol{y}_i^{(k)} = (\boldsymbol{y}_i^{\text{obs}^T}, \boldsymbol{y}_i^{\text{mis}(k)^T})^T$을 식 (2.4)에 대입하여 $\{e_{ij}^{(k)}\}$을 계산하되, 공통 공분산은

$$\boldsymbol{\Gamma}^{(k)} = \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{1}{n_i+1}\sum_{j=1}^{n_i+1}\left(\boldsymbol{y}_{j+}^{(k)} - \bar{\boldsymbol{y}}_i^{(k)}\right)^T\left(\boldsymbol{y}_{j+}^{(k)} - \bar{\boldsymbol{y}}_i^{(k)}\right)\right\}$$

와 같이 계산한다. 여기서 $\boldsymbol{y}_{j+}^{(k)} = (\boldsymbol{y}_j^{(k)^T}, \{\boldsymbol{y}_{\partial_i}^{(k)^T}\})^T$ 및 $\bar{\boldsymbol{y}}_i^{(k)} = \sum_{j=1}^{n_i+1}\boldsymbol{y}_{j+}^{(k)}/(n_i+1)$을 나타내며, n_i는 화소 i의 근방화소 개수이다.

그 다음 결측화소 대체값 $E[\boldsymbol{Y}_i^{\text{mis}}|\boldsymbol{y}^{\text{obs}}, \boldsymbol{e}^{(k)}, \boldsymbol{\theta}^{(k)}]$

$$\boldsymbol{y}_i^{\text{mis}(k+1)} = \boldsymbol{\mu}_i^{\text{mis}(k)} + \boldsymbol{\Sigma}_i^{\text{mo}(k)}\left(\boldsymbol{\Sigma}_i^{\text{oo}(k)}\right)^{-1}\left(\boldsymbol{y}_i^{\text{obs}} - \boldsymbol{\mu}_i^{\text{obs}(k)}\right), \quad i = 1, \ldots, n \quad (3.4)$$

과 $E[\boldsymbol{Y}_i\boldsymbol{Y}_i^T|\boldsymbol{y}^{\text{obs}}, \boldsymbol{e}^{(k)}, \boldsymbol{\theta}^{(k)}]$의 예측값

$$\boldsymbol{S}_i^{(k+1)} = \begin{pmatrix} \boldsymbol{S}_i^{\text{oo}(k+1)} & \boldsymbol{S}_i^{\text{om}(k+1)} \\ \boldsymbol{S}_i^{\text{mo}(k+1)} & \boldsymbol{S}_i^{\text{mm}(k+1)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{y}_i^{\text{obs}}\boldsymbol{y}_i^{\text{obs}^T} & \boldsymbol{y}_i^{\text{obs}}\boldsymbol{y}_i^{\text{mis}(k+1)^T} \\ \boldsymbol{y}_i^{\text{mis}(k+1)}\boldsymbol{y}_i^{\text{obs}^T} & \boldsymbol{M}_i^{(k)} + \boldsymbol{y}_i^{\text{mis}(k+1)}\boldsymbol{y}_i^{\text{mis}(k+1)^T} \end{pmatrix} \quad (3.5)$$

를 계산한다. 단,

$$M_i^{(k)} = \text{var}\left[Y_i^{\text{mis}} | y_i^{\text{obs}}, e^{(k)}, \theta^{(k)}\right] = \Sigma_i^{\text{mm}^{(k)}} - \Sigma_i^{\text{mo}^{(k)}} \left(\Sigma_i^{\text{oo}^{(k)}}\right)^{-1} \Sigma_i^{\text{om}^{(k)}} \tag{3.6}$$

을 나타낸다.

그리고 M-step에서는

$$\mu_i^{(k+1)} = \frac{\sum_{j \in \partial_i} e_{ij}^{(k)} y_i^{(k+1)}}{\sum_{j' \in \partial_i} e_{ij'}^{(k)}}, \quad i = 1, \ldots, n \tag{3.7}$$

및

$$\Sigma_i^{(k+1)} = \frac{\sum_{j \in \partial_i} e_{ij}^{(k)} S_i^{(k+1)}}{\sum_{j' \in \partial_i} e_{ij'}^{(k)}} - \mu_i^{(k+1)} \mu_i^{(k+1)^T}, \quad i = 1, \ldots, n \tag{3.8}$$

를 구한다. E-step과 M-step을 추정치가 충분히 수렴할 때까지 반복한다.

3.3. 가법 잡음과 정규화

만약 주어진 영상의 화소값 벡터 y_i에 영상 독립적인 백색잡음 $v_i \sim \mathcal{N}(0, \lambda I)$가 포함되어

$$y_i^* = y_i + v_i, \quad i = 1, \ldots, n$$

와 같이 오염되어 있다면, y_i^{obs}를 잡음으로부터 복원(restoration)하고 y_i^{mis}를 대체하는 문제가 된다. 여기서 $\lambda > 0$은 모든 화소의 색상 성분에 부가된 잡음의 공통 분산이다. 이 경우 EM 알고리즘은 E-step의 환경만 바뀌게 된다.

$E[Y_i^{\text{obs}^*}] = \mu_i$이고 $\text{var}(Y_i^{\text{obs}^*}) = \Sigma_i^{\text{oo}} + \lambda I$이므로, $\Sigma_i^{\text{oo}(k)}$ 대신 $\Sigma_i^{\text{oo}^*(k)} = \Sigma_i^{\text{oo}(k)} + \lambda I$를 사용하여, 식 (3.4)는

$$\begin{aligned} y_i^{\text{mis}(k+1)} &= E\left[Y_i^{\text{mis}} | y_i^{\text{obs}^*}, e^{(k)}, \theta^{(k)}\right] \\ &= \mu_i^{\text{mis}(k)} + \Sigma_i^{\text{mo}(k)} \left(\Sigma_i^{\text{oo}^*(k)}\right)^{-1} \left(y_i^{\text{obs}^*} - \mu_i^{\text{obs}(k)}\right), \quad i = 1, \ldots, n \end{aligned} \tag{3.9}$$

과 같이 바뀌고,

$$\begin{aligned} y_i^{\text{obs}(k+1)} &= E\left[Y_i^{\text{obs}} | y_i^{\text{obs}^*}, e^{(k)}, \theta^{(k)}\right] \\ &= \mu_i^{\text{obs}(k)} + \Sigma_i^{\text{oo}(k)} \left(\Sigma_i^{\text{oo}^*(k)}\right)^{-1} \left(y_i^{\text{obs}^*} - \mu_i^{\text{obs}(k)}\right), \quad i = 1, \ldots, n \end{aligned} \tag{3.10}$$

와 같이 복원하게 된다. 그리고 식 (3.5)은

$$\begin{aligned} S_i^{(k+1)} &= E\left[Y_i Y_i^T | y_i^{\text{obs}^*}, e^{(k)}, \theta^{(k)}\right] = \begin{pmatrix} S_i^{\text{oo}(k+1)} & S_i^{\text{om}(k+1)} \\ S_i^{\text{mo}(k+1)} & S_i^{\text{mm}(k+1)} \end{pmatrix} \\ &= \begin{pmatrix} O_i^{(k)} + y_i^{\text{obs}(k+1)} y_i^{\text{obs}(k+1)^T} & y_i^{\text{obs}(k+1)} y_i^{\text{mis}(k+1)^T} \\ y_i^{\text{mis}(k+1)} y_i^{\text{obs}(k+1)^T} & M_i^{(k)} + y_i^{\text{mis}(k+1)} y_i^{\text{mis}(k+1)^T} \end{pmatrix} \end{aligned} \tag{3.11}$$

단,

$$\boldsymbol{M}_i^{(k)} = \mathrm{var}\left[\boldsymbol{Y}_i^{\mathrm{mis}} | \boldsymbol{y}_i^{\mathrm{obs}*}, \boldsymbol{e}^{(k)}, \boldsymbol{\theta}^{(k)}\right] = \boldsymbol{\Sigma}_i^{\mathrm{mm}(k)} - \boldsymbol{\Sigma}_i^{\mathrm{mo}(k)} \left(\boldsymbol{\Sigma}_i^{\mathrm{oo}*(k)}\right)^{-1} \boldsymbol{\Sigma}_i^{\mathrm{om}(k)} \tag{3.12}$$

및

$$\boldsymbol{O}_i^{(k)} = \mathrm{var}\left[\boldsymbol{Y}_i^{\mathrm{obs}} | \boldsymbol{y}_i^{\mathrm{obs}*}, \boldsymbol{e}^{(k)}, \boldsymbol{\theta}^{(k)}\right] = \boldsymbol{\Sigma}_i^{\mathrm{oo}(k)} - \boldsymbol{\Sigma}_i^{\mathrm{oo}(k)} \left(\boldsymbol{\Sigma}_i^{\mathrm{oo}*(k)}\right)^{-1} \boldsymbol{\Sigma}_i^{\mathrm{oo}(k)} \tag{3.13}$$

와 같이 바꾸어 계산하게 된다. M-step에서는 평균과 분산을 각각 식 (3.7) 및 식 (3.8)과 동일하게 계산한다. 그리고 유사성 부여치 $\{e_{ij}^{(k)}\}$는 대체/복원값 $\boldsymbol{y}_i^{(k)} = (\boldsymbol{y}_i^{\mathrm{obs}(k)^T}, \boldsymbol{y}_i^{\mathrm{mis}(k)^T})^T$를 바탕으로 계산하면 된다.

본 논문에서 모수 λ는 추정하지 않고 조율상수(tuning constant)로 사용할 것이다. 물론 잡음이 포함되어 있지 않으면(즉 $\lambda = 0$이면) 식 (3.9)~(3.13)는 앞 소절의 원래의 공식으로 축소된다. 그러나 주어진 영상에 잡음이 포함되어 있지 않다고 판단하더라도 식 (3.4)와 (3.6) 대신 식 (3.9)와 (3.12)의 사용을 추천한다. 그 이유는 잡음이 없는 깨끗한 영상일수록 국소적으로 거의 같은 색상인 화소들이 모여있는 경우가 종종 있다. 이러한 경우 사용하는 컴퓨터의 유효숫자 처리 수준에서 $\boldsymbol{\Sigma}_i^{\mathrm{oo}(k)} \approx \mathbf{0}$이거나 0-정치 행렬이 될 수 있다. 알고리즘 수행 중 이런 상황이 발생하는 것을 방지하기 위해 어떤 작은 값의 λ를 사용하여 행렬 $\boldsymbol{\Sigma}_i^{\mathrm{oo}(k)}$를 정규화(regularization)할 필요가 있다. 본 연구의 다음 절 실험에서는 일률적으로 $\lambda = 5$(주어진 영상의 색상 성분이 [0, 255] 사이의 정수를 가지는 경우)를 사용하였다.

4. 영상 실험

실험에서는 147 × 113 크기의 색조영상(그림 4.1의 첫째 줄 첫 영상)의 R-, G-, B- 성분행렬에 각각 50%, 60%, 70%, 80%, 90%의 화소를 랜덤하게 결측 시켰다 (그림1의 첫 줄). 따라서 각 화소마다 (홍, 녹, 청) 성분 중 0, 1, 2개 혹은 3개 모두 결측될 수 있다. 예를 들어, 90% 성분 결측인 예의 경우 총 16,611개의 화소 중 99.89%(약 16,593개)의 화소에서 1개 이상의 원소들이 결측된 화소였고, 0.11%(약 18개)의 화소만이 결측성분이 없는 화소였는데, 1개 이상 성분 결측 화소들 중 1개 성분에서, 2개의 성분에서 그리고 3개 성분에서 모두에서 결측된 화소의 비율을 각각 2.78%(약 461개), 24.12%(약 4,006개) 및 72.99%(약 12,126개)였다. 원래 결측 성분이 있으면 화소의 색상을 표현할 수 없으나 결측 성분값은 0으로 간주하여 1개의 성분값 결측이면 노랑, 자주, 하늘색으로, 2개의 성분값 결측이면 적, 녹, 청색으로 그리고 3 성분 모두 결측인 화소는 검정색으로 나타내었다.

부여치 e_{ij}를 구하기 위한 함수의 매개변수는 $a = 1/2$, $b = 1$, $c = 0.01$로 하였고, 가법잡음 모형을 선택하고 잡음의 공통분산을 $\lambda = 5$로 하였다. 반복횟수는 70% 이하 결측영상의 경우는 30회 그리고 80% 이상의 경우는 50회 반복하였는데, 모든 경우 알고리즘은 단조적으로 안정되게 수렴하였다. 제안된 방법은 대중적으로 많이 이용되는 큐빅 스프라인 보간법(cubic spline interpolation; CSI)에 의한 대체 기법과 비교한다. CSI 기법의 구현은 Matlab에서 제공하는 함수 TriScatteredInterp( )를 이용하였다. 이 함수는 자체로 불연속을 반영하는 기능 즉 에지를 보존하는 기능을 가지고 있다. 아울러 본 연구에서는 CSI 기법의 결과를 제안된 방법의 초기 대체값 $\boldsymbol{y}^{(0)}$으로 이용하였다.

그림 4.1의 둘째 줄과 셋째 줄에 CSI 대체 영상과 제안된 기법에 의한 대체/복원 영상을 결측비율별로 각각 제시하였다. 두 기법 모두 높은 50% 이상의 높은 결측비율에도 불구하고 대체적으로 양호한 결과를 보여주고 있다. 그러나 객체의 면과 면이 만나는 부분 즉 에지 영역(특히 오른 쪽 머리카락 부분, 어깨 부분, 입 부분 및 눈동자 부분)에서 CSI 대체 영상은 주변과 상이한 색상으로 대체되어 있음을 발견할 수 있다. 이러한 현상은 결측률이 증가하면서 더욱 심해진다. 반면 제안된 방법의 결과는 CSI 대체

그림 4.1. 랜덤결측 비율에 따른 대체 결과(첫 줄: 원영상 및 50%, 60%, 70%, 80%, 90% 결측 영상, 둘째 줄: CSI에 의한 대체 영상, 셋째 줄: 제안된 방법에 의한 대체/복원된 영상)

그림 4.2. 80% 결측비율 영상에 대한 대체영상의 R, G, B-성분별 결과(첫 줄: 원영상 및 관측영상의 R, G, B-성분 , 둘째 줄: CSI 대체 영상 및 원영상과 차분, 셋째 줄: 제안된 방법에 의한 대체/복원된 영상 및 원영상과의 차분)

영상보다 훨씬 덜 한 결과를 보여주고 있다. 특별히 그림 4.2에는 80% 결측 영상에 대한 두 기법의 결과를 R, G, B 성분별로 나타내어 보았다. 두번째 줄과 세번째 줄에는 각각 CSI 기법의 결과 및 제안된

표 4.1. 예측오차

성분 결측율		50%	60%	70%	80%	90%
예측오차 $\sqrt{\mathrm{MSE}}$	CSI 기법	31.31	34.73	39.36	45.15	57.49
	제안된 방법	11.73	15.96	21.74	27.46	39.92
	상대비	2.67	2.18	1.81	1.64	1.44

방법의 결과와 원영상과의 차이를 R, G, B 성분별로 보여주고 있는데, 에지가 아닌 영역에서는 차이가 거의 없는 반면 에지 부분에서 CSI 기법에서의 차이가 제안된 방법보다 특히 두드러지게 나타나 있음을 알 수 있다.

표 4.1에 각 결측비율에 따른 결측화소의 성분값 당 예측오차($\sqrt{\mathrm{MSE}}$)을 수록하였다. 단,

$$\mathrm{MSE} = \frac{\sum_{i \in \mathcal{M}} \left(y_i - \hat{y}_i^{\mathrm{mis}}\right)^2}{m-1}$$

이며, $\mathcal{M}$은 모든 결측원소의 집합이며, m은 $\mathcal{M}$의 원소 개수를 나타낸다. 화소벡터의 성분값에 대한 예측오차가 색상의 차이를 판단하는 기술통계로서 적절한지는 의구심이 들지만, 아무튼 제안된 방법이 CSI 기법의 그것보다는 모든 결측비율에 걸쳐 낮게 평가되었다. 그러나 결측비율이 커지면서 그 상대비은 점차 1로 감소하는데, 이는 어쩌면 당연한 것으로서 결측비율이 90% 이상 높아지면 사용할 관측자료가 거의 없기 때문에 어떠한 대체 기법이라도 원영상을 제대로 복구할 수는 없음을 반영하는 것이라 하겠다.

참고로 본 실험에 사용된 언어와 기종은 Matlab (v.13)과 Intel Pentium(R), Dual Core(2.5GHz)이다.

5. 결론

본 논문에서는 색조 영상에서 결측화소값 성분의 랜덤결측값을 대체하는 문제를 다루었다. 본 논문의 요체는 주어진 화소와 근방화소들 사이의 유사성을 측정하여 근방화소들의 공분산 (행렬)에 부여치를 부과함으로써 동질적 주변화소만으로 대체값을 결정하도록 한다는데 있다. 영상 모형으로서 MRF의 국소성을 부여치로 '척도화한 공분산'을 가지는 정규분포의 결합으로 정의하고, 근사 E-step을 가진 EM 알고리즘으로 결측화소의 성분값을 대체하였다. 그 결과 매우 높은 결측비율에서도 매우 양호한 대체결과를 보여주었다.

다만, 근사 E-step을 가진 EM 알고리즘을 사용함으로써 그 수렴점이 상정한 모형의 우도와는 약간 다른 어떤 목적함수에 도달할 것이지만, 본 연구에서는 그 목적함수가 무엇인지는 확인하지 못하였다. 이 점은 추후 연구에서 명확히 밝힐 것이다.

참고문헌

김승구 (2009). 마코프 랜덤 필드 하에서 정규혼합모형에 의한 다중 결측값 대체기법: 색조영상 결측 화소값 대체에 응용, <한국통계학회논문집>, **16**, 914–925.

Besag, J. (1975). Statistical analysis of non-lattice data, *The Statistician*, **24**, 179–195.

Blanchet, J. and Vignes, M. (2009). A model-based approach to gene clustering with missing observation reconstruction in a Markov random field framework, *Journal of Computational Biology*, **16**, 475–486.

Dass, S. C. and Nair, V. N. (2003). Edge detection, spatial smoothing, and image reconstruction with partially observed multivariate data, *Journal of the American Statistical Association*, **98**, 77–89.

Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion), *Journal of Royal Statistical Society B*, **39**, 1–38.

Kim, D., Lee, Y. and Oh, H. S. (2006). Hierarchical likelihood-based wavelet method for denoising signals with missing data, *IEEE Signal Processing Letters*, **13**, 361–364.

Ogawa, T., Haseyama, M. and Kitajima, H. (2006). Restoration of missing intensity of still images by using optical flows, *System and Computers in Japan*, **37**, 1786–1795.

Qian, W. and Titterington, D. M. (1992). Stochastic relaxations and EM algorithms for Markov random fields, *Journal of Statistical Computation and Simulation*, **40**, 55–69.

Wei, G. C. G. and Tanner, M. A. (1990). A Monte Calro implementation of the EM algorithm and the poor man's data augmentation algorithms, *Journal of the American Statistical Association*, **85**, 699–704.

An EM Algorithm-Based Approach for Imputation of Pixel Values in Color Image

Seung-Gu Kim[1]

[1]Department of Computer & Data Information, Sangji University

(Received February 2010; accepted March 2010)

Abstract

In this paper, a frequentistic approach to impute the values of R, G, B-components in random missing pixels of color image is provided. Under assumption that the given image is a realization of Gaussian Markov random field, its model is designed such that each neighbor pixel values for a given pixel follows (independently) the normal distribution with covariance matrix scaled by an evaluates of the similarity between two pixel values, so that the imputation is not to be affected by the neighbors with different color. An approximate EM-based algorithm maximizing the underlying likelihood is implemented to estimate the parameters and to impute the missing pixel values. Some experiments are presented to show its effectiveness through performance comparison with a popular interpolation method.

Keywords: Imputation, random missing pixel, color image, Gaussian Markov random field, EM algorithm.

This research was supported by the Sangji University Research Fund 2009.

[1]Professor, Department of Computer & Data Information, Sangji University, 1 WooSan-Dong, KangWon-Do, Wonju 220-702, Korea. E-mail: sgukim@sangji.ac.kr