

퍼지적분을 이용한 영상품질의 객관적이고 정량적 평가: 팬텀 연구

가톨릭대학교 의과대학 의공학교실

김성현 · 서태석 · 최보영 · 이형구

물리적 평가(physical evaluation)가 영상품질의 객관화와 정량화를 위한 토대를 제공함에도 불구하고, 부정확하고 가변적인 특성을 지닌 주관적 평가(subjective evaluation)가 영상평가에 중요한 역할을 하게 된다. 본 연구에서는 디지털 방사선 영상의 물리적 평가와 주관적 평가의 단점을 상호 보완하고 객관적 정량화를 위한 새로운 방법을 제안하고자 한다. 임상에 사용되고 있는 4대의 디지털 방사선 영상 촬영장치로부터 동일한 임상조건에서 흉부 팬텀 영상을 획득하였다. 물리적 영상평가를 위하여 디지털 흉부 팬텀 내에서 3개의 영역(폐, 심장, 그리고 복부)에 존재하는 CNR (contrast-to-noise ratio)를 측정하였고 분할(segmentation)과 정합(registration)등 다양한 영상처리기술이 적용되었다. 주관적 평가는 5명의 관찰자에 의한 저 대조도 물체의 식별 정도를 점수화 하였다. 두 평가의 특성을 보완 및 결합하고자 퍼지적분 이론이 도입되었다. 4대의 시스템으로부터의 평가결과가 비교되었으며, 물리적 평가와 주관적 평가가 항상 비례하지 않음을 보였다. 물리적 평가에서는 높은 점수를 보였던 시스템이 주관적 평가에서는 상대적으로 낮은 평가를 보였다. 본 연구에서 제안한 퍼지적분에 의한 영상평가의 정량화는 물리적 평가와 주관적 평가를 모두 포함하는 총체적인 평가 방법이며, 다양한 영상품질 평가에 유용할 것이라 사료된다.

중심단어: 디지털 방사선 영상촬영시스템, 물리적 평가, 주관적 평가

서 론

100여년 전 뢰트겐이 x-ray를 발견한(1895) 이후 필름/스크린 방식이 주를 이루면서, 1982년 CR (computed radiography)의 등장으로 디지털 시대가 도래하였다. 이후 CCD (charge-coupled device)와 발광물질(scintillator)을 결합한 시스템들이 개발 되었다. 최근에 높은 DQE (detective quantum efficiency)³⁾의 특성을 가진 직/간접 방식의 평판형 디텍터(FPD: flat-panel detector)가 등장함으로써 새로운 디지털 시대로 접어 들었다.¹⁻²⁾

디지털 영상시스템에서, 영상정보는 두 단계에 의해 관찰자에게 전달된다. 첫번째 단계는 디텍터로부터 영상 형성 및 획득이며, 두번째 단계는 영상정보의 처리와 표시(display)이다. 첫번째 단계에서의 평가는 디텍터의 물리적 관점에서 묘사 되어질 수 있으며, 두번째 단계에서는 관찰

자의 인지도 관점에서 묘사되어 질 수 있다. 디텍터의 물리적 특성과 품질은 MTF (Module Transfer Function), NPS (Noise Power Spectrum), and DQE (Detective Quantum Efficiency) 등에 의해 특징 지어 진다. 또한 저 해상도의 탐지능 연구에 의해서도 평가되어지기도 한다. 의료분야에서 임상적 상황을 고려하면서 영상품질을 정량적으로 평가하는 것은 어려운 일이다. 임상을 고려한 유용한 영상평가는 ROC (Receive Operating Characteristic) 측정이나 대조도-상세 다이어그램(contrast-detail diagram)과 같은 주관적 평가에 의한 영상 평가법들이 널리 사용되고 있다.³⁾ 물리적 평가는 객관화와 정량화를 위한 중요한 토대를 제공하지만, 인간의 시각계에 의한 평가를 고려하지 않는다. 주관적 평가는 가변적이고 부정확한 특징을 가지지만 영상평가에 중대한 역할을 한다. 영상표시장치 또한 인간의 시각계를 통한 영상 인식에 중대한 영향을 미치기 때문에, 이들의 특성을 이용한 영상처리연구가 이루어졌다.^{4,7)}

본 연구에서는 디지털 방사선 영상의 물리적 평가와 주관적 평가의 장/단점을 상호 보완하고 객관적 정량화를 위한 새로운 방법을 제안하고자 한다. 물리적 평가를 위해 디지털 흉부 팬텀영상의 3개의 영역에서(폐, 심장 그리고 복부)저 대조도 물체의 CNR을 정확하고 신속하게 측정하기

이 논문은 2008년 10월 21일 접수하여 2008년 12월 18일 채택되었음.
 책임저자 : 이형구, (137-040) 서울시 서초구 반포동 505번지
 가톨릭대학교 의과대학 의공학교실
 Tel: 02)590-2415, Fax: 02)532-1779
 E-mail: hkleee@catholic.ac.kr

위한 영상분석 알고리즘이 개발되었다. 그리고 해당 영역에서 관찰자들의 저 대조도 물체의 식별 정도를 점수화 하였다. 영상품질의 주관적 평가를 표현하기 위해, 적당한 방법으로 알려진 퍼지측도 이론이 사용되었으며, 마지막으로 퍼지적분에 의해 두 평가 결과가 결합되었다. 퍼지적분은 어떤 대상이 여러 항목에 대해서 평가가 되고 이들 평가치를 종합하는데 이용될 수 있는 것으로 알려져 있다.

연구결과는 동일한 실험조건에서 동일한 제품의 디텍터를 사용하는 4대의 상업용 영상처리장치에서 상호 비교 되었다.

재료 및 방법

1. 영상획득

연구를 위하여 아모퍼스-실리콘(PIXIUM 4600, TRIXELL, France) 디텍터를 사용하는 4대의 상업용 영상처리장치들(시스템_A, 시스템_B, 시스템_C 그리고 시스템_D)이 채택되었다. 이들 디텍터의 물리적 특성은 Chaussat⁸⁾ 등에 의해 보고 된 바 있으며, 탈륨(Tl)과 합금된 Cesium-Iodide (CsI)가 500 μm 의 층을 이룬다. 이것은 엑스선을 가시광선으로 전환시키는 역할을 하며, 아모퍼스-실리콘 층이 빛을 전기적 신호로 전환시키게 된다. 영상처리 장치를 통하여 픽셀의 크기가 143 μm (3,121×3,121 matrix size), 43×43 cm^2 크기를 가진 12 비트 영상을 창출해 낸다. 연구의 목적이 시스템들의 영상의 비교 평가가 아니라 영상 평가를 위한 새로운 방법을 제안하는 것이기 때문에 이들 장치들의 자세한 내용은 언급하지 않는다. 이들 시스템들의 큰 차이점은 영상의 품질을 담당하는 영상처리장치가 다르다는 것이다. 동일한 디텍터를 사용함에도 불구하고 영상처리기술에 따라서 영상의 품질은 서로 다른 특징들을 보인다. 즉, 방사선의 산란선에 의해 발생한 노이즈를 제거하거나, 노이즈의 증강 없이 대조도와 경계선을 시각적으로 향상시키는 기술 등은 영상의 품질에 상당한 영향을 미친다. 본 연구에서 영상처리장치의 비교는 영상처리기술의 비교라 할 수 있을 것이다.

4대의 영상시스템으로부터의 영상획득은 모두 동일한 조건에서 이루어 졌다. 이들 시스템은 동일한 디텍터에 동일한 사양의 고정형 그리드(12 : 1, 3.9 lines/mm), 그리고 0.6 mm 초점을 사용하였다. 시험을 위한 촬영거리는 소스와 디텍터간 거리를 180 cm으로 하였다. 엑스레이 촬영조건을 설정하기 위해 먼저 21 mm 알루미늄 판을 엑스레이 소스의 가장 가까운 부분에 부착하여 125 kVp 조건에서 입

사선량이 5 μGy 가 되도록 하였다. 선량측정은 20×6-60 이온전리함을 내장한 2026C모델(Radcal, CA, USA)을 사용하였으며, 후방 산란을 최소화하기 위해 디텍터 앞 40 cm 거리에서 선량을 측정하였다. 제곱 반비례 법칙에 의해 소스로부터 180 cm의 거리에서의 입사선량이 도출되었다.

본 연구를 위하여 Fig. 1과 같은 흉부 팬텀(Nuclear Associates Model 07-646, Nuclear Associates, Carle Place, NY)이 사용되었다. 간략하게 설명하자면, 흉부 팬텀 영상의 폐, 심장, 그리고 복부에 각각 다른 지름과 두께를 가진 구리로 구성된 25개의 원모양의 저 대조도 물체들이 일정한 배열로 존재한다. 3영역에서 테스트 물체들이 존재하는 부분을 제외하고는 저 선량에도 시각적으로 뚜렷하게 보이는 스테인레스 와이어로 망이 구성되어 있다. 21 mm 알루미늄 판을 제거 후 이미 설정된 조건하에서 4시스템으로부터 흉부 팬텀 영상을 획득하였다. 각 시스템의 영상처리장치를 통하여 영상저장전송장치(Picture Archiving and Communication System) 네트워크로 12비트 팬텀 영상이 전송되었다.

2. 영상품질 평가

저 대조도의 정확한 CNR을 측정하기 위해 영상분석 알고리즘이 개발되었다. 흉부 팬텀을 이용한 저 대조도의 정량적 분석을 위한 방법은 이미 보고된 바가 있지만 고 선량을 요구한다.⁹⁾ 문헌 9)에서는 정량적이고 객관적인 영상

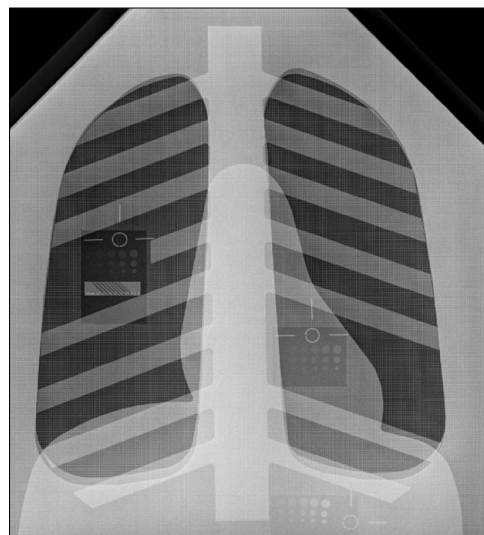


Fig. 1. Image of a QC chest phantom used for the physical and subjective evaluation. Three representative ROIs were selected for the evaluation (Lung, Heart, and Diaphragm). There are 25 low contrast objects in each region.

평가를 위하여 영상의 정합 알고리즘을 소개하며, SNR의 측정에 있어서 객관성을 강조 하였다. 그러나 저 대조도를 가진 물체들의 경계선을 추출하여, 물리적 측정이 이루어지기 때문에 고 선량이 요구된다. 이는 저 선량을 사용하여 환자의 피폭량을 감소시켜야 하는 임상에서는 실용적이지 못한 것으로 사료된다. 본 연구에서는 팬텀 구조물들을 기준표시점(fiducial marker)으로 활용하여 물체의 중심위치를 자동으로 파악하도록 알고리즘¹⁰⁾을 구현 하였으며, 이를 통하여 저 선량에서도 저 대조도 물체의 위치를 파악하여 CNR측정이 가능하도록 하였다. 주관적 평가에서는 퍼지측도가 적용되었으며, 퍼지적분에 의해 물리적 평가와 통합되었다. 획득한 4개의 흉부 팬텀 영상이 사용되었다. Fig. 2에 영상품질을 평가하기 위한 전체적인 순서도를 나타내었다

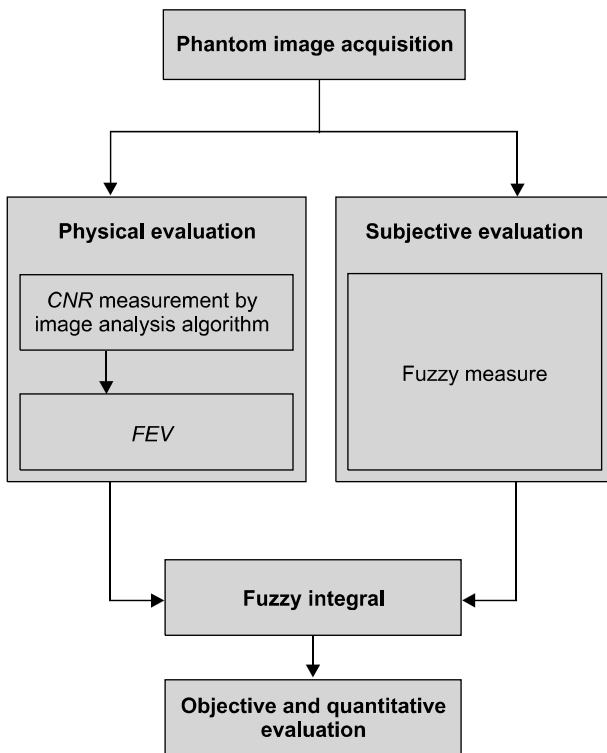


Fig. 2. The workflow for image evaluation. Chest phantom images were acquired for physical and subjective evaluations. In physical evaluation, CNRs of nine most visible objects in each chest region were derived from image analysis algorithm. 25 objects in each region were detected and scored by human observers. These were converted to FEV and Fuzzy measure for application to fuzzy integration. Finally, fuzzy integration of physical and subjective properties was performed for objective and quantitative evaluation.

3. 물리적 평가

1) **영상 분할과 템플레이터:** 저 대조도 물체의 절대적 위치를 파악하기 위한 물체 위치 템플레이터를 획득하기 위해, 먼저 저 대조도 물체가 가장 잘 보이는 10장의 팬텀 영상을 획득하여 평균화하여 노이즈가 최소화된 영상을 얻었다. 그레이 값의 분포를 확장시키고 영상의 보이지 않는 부분을 더 잘 보일 수 있도록 CLAHE (contrast-limited adaptive histogram equalization) 기법이 적용되었다. 히스토그램의 clipping height는 2.15로 설정 하였다. Bilateral 필터를 적용한 후, component labeling 기법¹⁰⁻¹¹⁾을 이용하여 분할(segmentation)이 이루어졌다. 팬텀의 각 영역(폐, 심장, 복부)에서 가장 잘 보이는 9개의 저 대조도 물체의 외곽이 추출되었다. 추출된 원 모양의 영역은 foreground로 정의 되었다.

2) **정합:** foreground의 중심이 입력영상의 해당하는 물체의 중심과 일치하도록 foreground 템플레이터를 입력영상과 정합하였다. 흉부 팬텀의 납 구조물들은 저 선량에도 불구하고 좋은 선명도를 유지한다. 이러한 특징 때문에 주변의 구조물들은 물체의 중심을 찾기 위한 기준표시점으로 사용되었다. 영상의 정합은 foreground 템플레이터의 이동, 회전, 그리고 크기 조절에 의해 구현 되었다. Fig. 3은 템플레

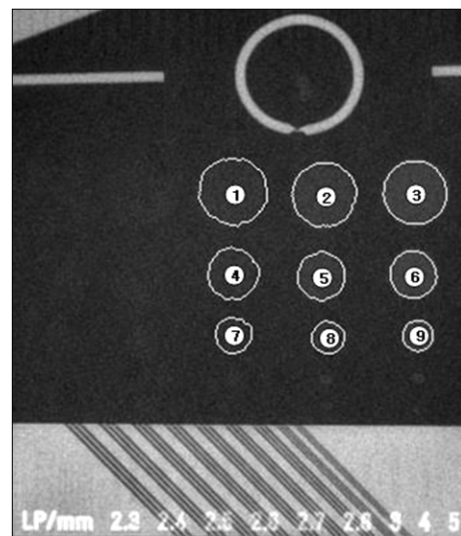


Fig. 3. Low-contrast objects in the lung region. The low-contrast objects in each column have the same low contrast level, and those in each row have the same size. The foreground template is composed of the boundary of the nine most visible objects. The foreground template is superimposed on the input image with the low-contrast objects by translation, rotation, and scaling.

이터와 입력영상과의 정합결과를 보여준다. foreground와 background에 해당하는 영역의 픽셀 값들은 흉부 팬텀 영상의 3영역에서의 저 해상도 물체의 CNR을 계산하기 위해 사용되었다.

3) CNR 측정: CNR을 평가하기 위해 Fig. 4에서 처럼 각 물체에 대해 1개의 foreground 영역과 주위의 4개의 background 영역이 정의 되었다. Background 영역은 11픽셀의 반경을 가진 원이 되도록 하였다. 실제 foreground는 완전한 원 모양은 아니었지만 본 연구의 결과에 영향을 미치지 않는 않았다. 흐려짐을 최소화 하기 위해 foreground 외곽으로부터 1픽셀(pixel) 떨어진 영역의 픽셀들이 사용되었다. 신호는 foreground와 background의 평균 픽셀 값의 차이를 계산함으로써 획득 되었다. 노이즈는 5개의 영역에서 픽셀 값들의 평균 표준편차로 결정 되었다. CNR은 다음의 수식에 의해 정의 되었다.

$$CNR = \frac{|E_{FG} - E_{BG}|}{\sqrt{(V_{FG} + V_{BG})/2}} \quad (1)$$

여기서 E_{FG} 와 E_{BG} 는 foreground와 background 영역의 평균 픽셀 값들이며, V_{FG} 와 V_{BG} 는 이들 영역에서의 분산 값들이다. 흉부영역의 각 영역 x_i (x_1 : 폐, x_2 : 심장, x_3 : 복부)에서 9개의 물체들의 CNR의 대표값을 결정하기 위해 칸텔에 의해 소개되었던 퍼지 기대값(FEV: Fuzzy Expected Value) 개념이 사용되었다.¹²⁾ 이것은 주어진 퍼지 집합에서 소속하는 정도의 대표값을 나타내는 것으로, 중간값이나 평균값

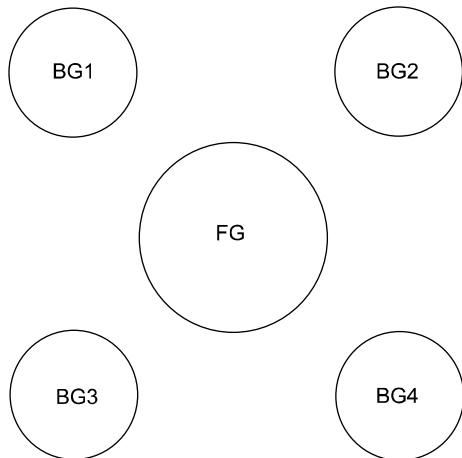


Fig. 4. One foreground (FG) region and four background (BG1, BG2, BG3, BG4) regions of the sampling template. FG is within a low-contrast object. The background region was selected to be a circle with a radius equal to 11 pixels.

개념 이상의 새로운 양(quantity)으로 묘사 되어 진다. CNR은 퍼지 기대값을 도출하기 위해 퍼지 영역으로 전환되었다. 9개의 저 대조도 물체의 집합 $A = \{\alpha_1, \alpha_2, \dots, \alpha_9\}$ 에 대하여 퍼지밀도 $f: CNR(A) \rightarrow [0,1]$ 가 정의되고, $\xi_T = \{\alpha | f(\alpha) \geq T\}$, $0 \leq T \leq 1$ 이며, μ 는 A의 부분집합의 퍼지측도이다.

Table 1. CRS and FEV in lung, heart, and diaphragm regions from four systems.

(a) CRS and FEV in lung region.

Object	System-A	System-B	System-C	System-D
1	2.467	1.160	1.424	1.578
2	3.789	2.363	2.332	2.847
3	5.320	3.596	3.454	4.191
4	1.843	1.409	1.533	1.349
5	3.625	2.340	2.258	2.711
6	5.544	3.665	3.374	4.355
7	1.436	1.244	1.350	1.089
8	3.551	2.158	2.457	2.381
9	5.304	3.577	3.507	4.205
FEV	0.708	0.257	0.462	0.487

(b) CRS and FEV in heart region.

Object	System-A	System-B	System-C	System-D
1	1.971	1.133	1.638	1.891
2	2.762	2.008	2.271	2.823
3	4.549	2.905	3.403	3.996
4	1.868	1.272	1.531	1.815
5	3.162	2.113	2.345	2.577
6	5.148	3.084	3.415	4.317
7	2.220	1.146	1.531	1.784
8	2.855	1.663	2.512	2.670
9	4.690	2.950	3.400	4.398
FEV	0.538	0.220	0.373	0.487

(c) CRS and FEV in diaphragm region.

Object	System-A	System-B	System-C	System-D
1	2.524	1.588	2.110	2.998
2	4.169	2.541	3.255	5.389
3	6.670	3.653	4.771	6.224
4	2.899	1.695	2.214	3.311
5	4.113	2.497	3.232	4.080
6	6.827	3.608	4.671	6.862
7	3.041	1.459	2.280	3.015
8	4.019	2.538	3.171	4.142
9	6.571	3.629	4.693	6.758
FEV	0.739	0.490	0.637	0.743

$$h(x_i) = FEV(x_i) = \text{Max}\{\text{Min}[T(x_i), \mu(x_i, \xi_T)]\} \quad (2)$$

여기서 $h(x_i)$ 은 각 영역 x_i 에서의 CNR들의 퍼지 대표값을 나타낸다. 자세한 내용은 문헌 12)를 통하여 확인할 수 있다. Table 1에 4대의 장치에서 획득한 흉부 팬텀 영상의 각 영역에서의 측정된 CNR와 퍼지 기대값을 나타내었다. 이 결과 값은 주관적 평가결과와 퍼지 적분에 의해 결합되어 진다.

4. 주관적 평가

주관적 평가는 관찰자에 의한 저 대조도 물체의 탐지능에 의해 평가 되었다. 물리적 평가에서 입력영상으로 사용되었던 4개의 흉부 팬텀 영상은 동일한 판독전용 모니터상에 표시 되었다. 관찰이 용이하게 동일한 휘도의 공간에서 이루어 졌다. 관찰자 들에게 영상과 관련하여 어떠한 정보도 주어 지지 않았으며 영상은 무작위로 제공 되었다. 2명의 방사선과 전문의와 3명의 방사선사로 구성된 5명의 관찰자에 의해 수행되었으며, 영상을 관찰하는 동안 영상의 밝기와 대조도 조절 혹은 관찰 거리, 시간과 같은 어떠한 제한 없이 저 대조도 물체를 식별 하도록 하였다. 흉부 팬텀 영상의 각 영역에서 25개의 저 대조도 물체의 식별 여부를 점수화 하여 평균값을 얻었으며, 최소한 50%이상 보이는 물체에 대해서만 유효가 부여 되었다. 관찰자 자체의 가변성에 의한 영향을 최소화 하기 위해서 3일 주기로 3번의 평가가 이루어 졌으며 그 평균값을 구하였다. 4대의 장치로부터 획득한 흉부 팬텀 영상의 각 영역에서의 결과 값을 Table 2에 나타내었다. 결과 값들은 퍼지적분에 적용하기 위해 퍼지측도로 전환 되어졌다.

5. 퍼지 적분을 이용한 통합평가

본 논문에서는 퍼지적도를 사용하여 수계노가 정의한 퍼지적분을 이용하였다.¹¹⁾ 수계노의 퍼지적분은 임의의 대상

을 평가한 후 이를 종합하는 성질이 있다. 특히 주관적인 판단이 개입되는 평가 문제에서 주요하게 활용되어 왔다. 퍼지측도는 인간이 가지는 주관적 평가나 판단의 측도로 사용되어 진다. 주관적 평가는 퍼지측도를 구하기 위해서 퍼지 밀도[0,1]로 전환되었다.

퍼지측도는 집합 $X(x_1, x_2, \dots, x_n)$ 에 대하여 3가지 조건, 연속성, 경계조건, 단조성을 만족시킨다. 본 연구에서는 수계노에 의해서 제안된 λ -퍼지측도를 적용하였다. 즉, x_1 과 x_2 가 공통점을 갖지 않을 때 $-1 < \lambda$ 에 대하여,

$$g_\lambda(x_1 \cup x_2) = g_\lambda(x_1) + g_\lambda(x_2) + \lambda \cdot g_\lambda(x_1) \cdot g_\lambda(x_2) \quad (3)$$

이때 $g_\lambda(X)=1$ 이고, $g_\lambda(\cdot)$ 는 퍼지측도의 조건을 만족한다. 여기서 $x_i (i=1,2, \dots, n)$ 는 이미 언급한 흉부영상에서의 영역들이며, λ 의 값에 따라서 다른 성질의 퍼지 측도가 된다. λ 의 값은 조건 $g(X)=1$ 로부터 계산되어 질 수 있다. 여기에서 $n=3$ 이다.

$$\lambda + 1 = \prod_{i=1}^n (1 + \lambda g(i)) \quad (4)$$

4영상시스템들의 주관적 평가를 나타내는 퍼지측도는 수식 (3)과 (4)에 따라서 도출 되었으며, 결과는 Table 3에 요약 되었다. Table 3에서 $\{x_1\}, \{x_2\}$, 그리고 $\{x_3\}$ 의 g_λ 는 이미 획득한 주관적 평가의 퍼지 밀도이며, 이들과 수식 4로부터 λ 를 도출한다. 보다 자세한 내용은 문헌 13)에서 확인할 수 있다. 수계노의 퍼지적분은 어떤 대상을 여러 항목에 대해서 평가할 때 이들 각 항목에 대한 평가치를 퍼지측도를 사용하여 종합하는 방법으로 사용될 수 있다. 전체 평가항목 X 에서

Table 3. Values of g_λ , x_1 -Lung, x_2 -Heart, x_3 -Diaphragm. Subjective image quality evaluation of the four different imaging systems. x_1, x_2 , and x_3 are the normalized subjective evaluations of the low-contrast detectability on the three regions (Lung, Heart, Diaphragm). λ was calculated using Eq. 3 to construct the Sugeno measure.

Subset	System-A	System-B	System-C	System-D
	0	0	0	0
{x1}	0.493330	0.466660	0.520000	0.586666
{x2}	0.533330	0.560000	0.653333	0.573333
{x3}	0.653330	0.613333	0.773333	0.760000
λ	-0.866296	-0.847474	-0.948693	-0.941352
{x1,x2}	0.798734	0.805193	0.851030	0.843371
{x1,x3}	0.867449	0.837434	0.911832	0.926949
{x2,x3}	0.884811	0.882254	0.947344	0.923155
{x1,x2,x3}	1	1	1	1

Table 2. Subjective averaged scores in lung, heart, and diaphragm regions from four systems.

System/Region	Lung	Heart	Diaphragm
System-A	12.325	2.665	16.325
System-B	11.65	14	15.325
System-C	14.65	14.325	19
System-D	13	16.325	19.325

의 물리적 평가 h 와 주관적 퍼지측도 g_λ 를 이용하여 종합평점을 구하였다. $x_i \in X (i=1, n)$ 에 대해 $h(x_i) \leq h(x_{i+1})$ 이라 하고, X 의 멱집합의 원소 E 에 대해 $E_i = \{x_k | k = i, \dots, n\}$ 이라 하자. 퍼지적분은 다음과 같다.

$$\int h(x) \circ g_\lambda = \text{Max}_{i=1}^n \text{Min} [\text{Min} h(x_i), g_\lambda(E_i)] \quad (5)$$

여기서 $n=3$ 이며 $\text{Min} [\text{Min} h(x_i), g_\lambda(E_i)]$ 에서 x_i 는 E_i 의 원소이고 가장 작은 $h(\cdot)$ 를 가진다. 평가항목 중 가장 보수적인 평가치 h 와 평가항목 E 의 퍼지측도 중에서 가장 작은 것을 선택한다. 이것은 가장 보수적인 평가치를 가짐과 동시에, 물리적 평가 h 가 주관적 평가 g_λ 보다 클 수 없다는 철학이 깔려 있다. 한편 적분결과를 $\text{Max}_{i=1}^n \text{Min} [\text{Min} h(x_i), g_\lambda(E_i)]$ 로 함으로써 여러 가지 가능한 E 중에서 가장 큰 값을 취하여 전체 평가치를 종합하고 있다. 이 부분에서는 긍정적인(유

리한) 항목을 부각시켜 낙관적인 평가를 하는 측면이 있다.

결 과

영상품질의 물리적 분석을 위해 영상분석 알고리즘이 개발되었고, 결과로 나온 물체 위치 템플레이터는 획득된 입력영상과 정합 되었다. 흉부 팬텀의 3개의 영역에서 각각 9개의 가장 잘 보이는 저 대조도 물체들이 추출 되었으며, 물체들의 CNR이 계산되어 졌다. Table 1에 나타나 있는 각 영역에서의 CNR은 퍼지 기대값으로 전환되었다. 퍼지 기대값외에 중간값이나 평균값으로 대표값을 고려하는 것도 의미 있을 것이다. 정합 알고리즘의 정확도는 절대적 위치와 비교하여 1픽셀의 오차를 벗어나지 않았다. 주관적 평가에서는 팬텀 영상의 각 영역에서 25개의 저 대조도 물체들을 5명의 관찰자가 3일 주기로 3번 평가 하였다. 평가 결

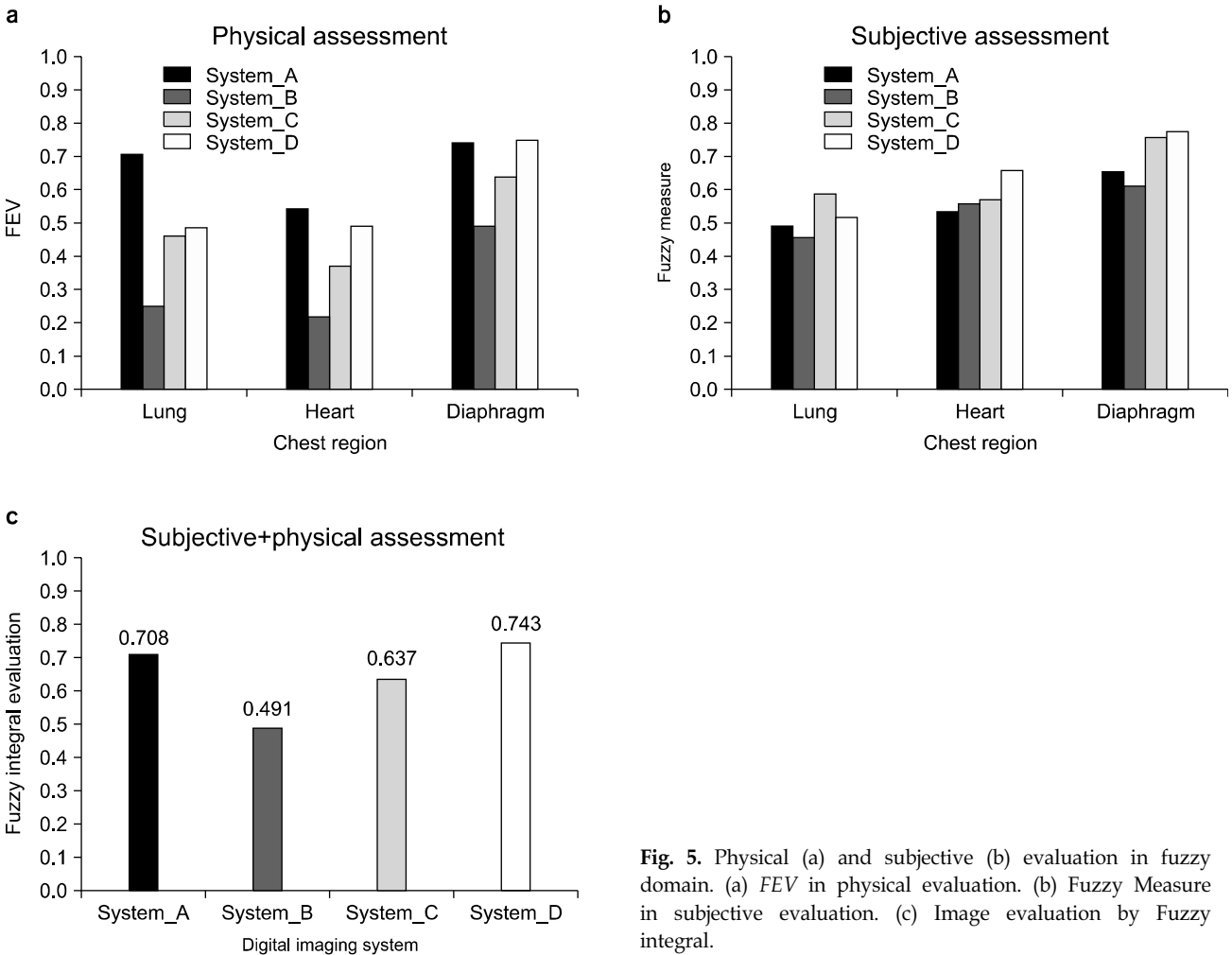


Fig. 5. Physical (a) and subjective (b) evaluation in fuzzy domain. (a) FEV in physical evaluation. (b) Fuzzy Measure in subjective evaluation. (c) Image evaluation by Fuzzy integral.

과의 평균값(Table 2)을 구하여 퍼지측도로 전환하였다. 본 연구에서 제안한 평가 방법을 적용하기 위해 4대의 영상 시스템으로부터 흉부 팬텀 영상을 획득하여 상호 비교 하였다. Fig. 5은 시스템_A, 시스템_B, 시스템_C 그리고 시스템_D에서 퍼지 영역에서의 물리적/주관적 평가 결과이다. Fig. 5a는 각 영역에서의 물리적 평가결과를 보여준다. CNR의 퍼지 기대값은 시스템_A에서 상대적으로 좋은 결과를 보여준다. 주관적 평가결과를 보여주는 Fig. 5b에서는 시스템_D에서 상대적으로 좋은 결과가 나왔음을 보여준다. Fig. 5c에서는 퍼지적분에 의해 통합된 결과이며, [0,1]의 퍼지 영역에서 시스템_D의 영상품질이 시스템_A보다 근소한 차이로 높음을 보여준다. Fig. 5a에서 물리적 측정에 의한 결과 값은 시스템마다 뚜렷한 차이를 보이는 것을 확인할 수 있다. 그러나 Fig. 5b에서 처럼 주관적 평가에 의한 결과는 시스템간의 차이가 감소함을 확인할 수 있다. 특히 시스템_A와 시스템_B는 (a)에서 뚜렷한 차이를 보였지만 (b)에서는 비슷한 평가치를 보인다. 이미 언급했던 것처럼 본 연구에 채택된 4대의 시스템들은 동일한 물리적 특성을 가지는 디텍터를 사용하고 있다. 따라서 본 연구결과로부터 확인할 수 있는 주요한 특징은 영상의 품질은 영상처리기술에 따라 물리적 평가와 주관적 평가에서 상이한 특징을 보인다는 것이다.

언급했던 것처럼 두 평가법의 장단점을 상호 보완하기 위해 평가들을 종합하는 특성을 지닌 퍼지적분 이론을 도입하였다. 본 연구의 결과에 대한 신뢰성과 유용성을 검증하는 것이 향후 지속적으로 수행되어야 할 과제라고 사료된다. 그러나 한편으로는 본 연구에서 도입한 퍼지적분 이론의 신뢰성이 문헌들로부터 입증되었다는 점에서 본 연구결과에 신뢰성을 두고 있다.

고찰 및 결론

디지털 방사선 촬영술과 더불어 디지털 영상처리기술 또한 발전하고 있으며, 그 역할이 더욱 더 중요해 지고 있다. 디지털 방사선 영상 시스템들은 해부학적인 혹은 병리학적인 특징을 향상 시키거나, 산란선에 의한 노이즈를 제거함으로써 임상적으로 유용한 양질의 영상을 제공한다. 이러한 방사선 영상의 임상적 유용성 혹은 품질 평가를 위해 물리적 평가와 주관적 평가가 사용 되어 진다. 본 연구에서는 흔히 사용되는 흉부 팬텀 영상을 이용하여, 영상의 물리적 평가와 주관적 평가를 시행 하였다. 일반적으로 물리적 평가가 높은 영상이 주관적 평가에서도 좋은 결과가 나올

것으로 예측한다. 그러나 본 연구의 평가결과에서는 일치하지 않음을 확인할 수 있었다. 그러나 물리적 측정과 주관적 측정 사이에 어떠한 논리적인 상관관계는 파악할 수 없었다. 관찰자의 가변적이고 부정확한 특성에도 불구하고, 주관적 평가가 중요한 역할을 해야 한다면, 물리적 평가만으로 영상품질을 결정하는 것은 한계가 존재하는 것으로 사료된다. 또한 유사한 연구가 있거나 주관적 평가와 물리적 평가의 비교에 그쳤다.¹⁴⁾ 따라서 본 연구의 가치를 찾는다면 주관적 평가와 물리적 평가의 장단점을 보완하기 위해서 퍼지적분 이론을 도입하여, 보다 객관적이고 정량적인 영상평가를 시도한 것이라 하겠다. 그러나 제안된 평가 방법은 이제 초기 모델로서 앞으로 계속해서 연구 되어 지고 정제 되어져야 할 것으로 사료된다.

감사의 글

본 연구를 위해 많은 도움을 주신 가톨릭의과대학 성모병원 방사선과, 한양대학교 의과대학 방사선과 선생님들 그리고 피닉스비전 연구원님들에게 감사의 말씀을 전합니다.

참 고 문 헌

1. Neitzel U, Maack I, Gunther-Kohfahl S: Image quality of a digital chest radiography system based on a selenium detector. *Med Phys* 21:509-516 (1994)
2. Siewerdsen JH, Antonuk LE, El-Mohri Y, Yorkston J, Huang W, Cunningham IA: Signal, noise power spectrum, and detective quantum efficiency of indirect-detection flat-panel imagers for diagnostic radiology. *Med Phys* 25:614-628 (1998)
3. Rong XJ, Shaw CC, Liu X, Lemacks MR: Comparison of an amorphous silicon/cesium iodide flat-panel digital chest radiography system with screen/film and computed radiography systems - A contrast-detail phantom study. *Med Phys* 28: 2328-2335 (2001)
4. Nawfel RD, Chan KH, Wagenaar DJ, et al: Evaluation of Video Grayscale Display. *Med Phys* 19:561-567 (1992)
5. Granrath DJ: The role of human visual models in image processing. *Proce IEEE* 69:552-561 (1981)
6. Ji TL, Sundareshan MK, Roehrig H: Adaptive Image Contrast Enhancement Based on Human Visual Properties. *IEEE Trans. Med Imag*, vol 13:573-586 (1994)
7. Suzuki J, Furukawa I, Ono S, et al: Contrast Mapping and Evaluation for electronic X-Ray images on CRT display monitor. *IEEE Trans. Med Image* 16:772-784 (1997)
8. Chaussat C, Chabbal J, Ducourant T, Spinnler V, Vieux G, Neyret R: New Csi / a-Si 17"x17" flat panel detector provides superior detectivity and immediate direct digital output for general radiography system. *SPIE Medical Imaging* 3336:33-44

- (1998)
9. Alexander LCK, Larry JF, Lawrence HL: Automatic quantitative low-contrast analysis of digital chest phantom radiographs. *Med Phys* 30:312-320 (2003)
 10. 김성현, 서태석, 윤세철 외: QA용 두부 팬텀을 이용한 Hitchcock CT 및 혈관조영 정위적 시스템에 대한 정확도 평가, *의학물리*, 9:1-9 (1998)
 11. Ballard D. Brown C: *Computer Vision*, Prentice-Hall, New York (1993), pp. 277-279
 12. Rafael C. Gonzalez, Richard E. Woods: *Digital Image processing*. Addison Wesley, New York (1992) pp. 341-343
 13. 이광현, 오길복: 퍼지이론 및 응용, 흥릉과학출판사, 서울 (1997), pp. 9_3-9_35
 14. Sandorg M. Tingberg A, Ullman G. Dance DR, Carlsson GA: Comparison of clinical and physical measures of image quality in chest and pelvis computed radiography at different tube volages. *Med Phys* 33:4169-4175 (2003)

Objective and Quantitative Evaluation of Image Quality Using Fuzzy Integral: Phantom Study

Sung-Hyun Kim, Tae-Suk Suh, Bo-Young Choe, Hyoung-Koo Lee

Department of Biomedical Engineering, College of Medicine, The Catholic University of Korea, Seoul, Korea

Physical evaluations provide the basis for an objective and quantitative analysis of the image quality. Nonetheless, there are limitations in using physical evaluations to judge the utility of the image quality if the observer's subjectivity plays a key role despite its imprecise and variable nature. This study proposes a new method for objective and quantitative evaluation of image quality to compensate for the demerits of both physical and subjective image quality and combine the merits of them. The images of chest phantom were acquired from four digital radiography systems on clinic sites. The physical image quality was derived from an image analysis algorithm in terms of the contrast-to-noise ratio (CNR) of the low-contrast objects in three regions (lung, heart, and diaphragm) of a digital chest phantom radiograph. For image analysis, various image processing techniques were used such as segmentation, and registration, etc. The subjective image quality was assessed by the ability of the human observer to detect low-contrast objects. Fuzzy integral was used to integrate them. The findings of this study showed that the physical evaluation did not agree with the subjective evaluation. The system with the better performance in physical measurement showed the worse result in subjective evaluation compared to the other system. The proposed protocol is an integral evaluation method of image quality, which includes the properties of both physical and subjective measurement. It may be used as a useful tool in image evaluation of various modalities.

Key Words: Digital radiography system, Physical evaluation, Subjective evaluation