

입체효과 최적화를 위한 사용자 보조 소프트 컴퓨팅 기법

User Assistant Soft Computing Method for 3D Effect Optimization

최우경, 김종수, 하상형, 김성현*, 전홍태
중앙대학교 전자전기공학부
*동원대학 디지털정보전자과

Woo-Kyung Choi, Jong-Soo Kim, Sang-Hyung Ha, *Sung-Hyun Kim, Seong Hong-Tae Jeon
School of Electrical and Electronic Engineering, Chung-Ang University
*Dept. of Digital Information Electronics, Tong-Won College
E-mail : chwk001@wm.cau.ac.kr

요 약

본 논문에서는 신경망 학습을 위한 데이터 획득시 생길 수 있는 오차를 줄이기 위해 획득 데이터에 대한 전처리 과정을 퍼지로서 구현하는 알고리즘을 제안하였다. 신경망은 주어진 정보를 이용하여 학습을 가능하게 함으로써 시스템의 특징을 추출하는데 매우 우수한 능력을 발휘하고 있다. 그러나 이는 학습에 사용하는 데이터에 오차가 포함되지 않는다는 점을 전제로 하고 있다. 그런데 데이터 획득과정이 인간의 주관적 판단에 의해 수작업으로 이루어지는 경우 학습 데이터는 오차가 존재할 수 있다. 학습 데이터의 오차를 줄이기 위해 초기에 획득된 데이터를 분석하고 추가적인 후보 데이터를 선정하여 데이터 획득 과정에서 큰 영향을 미치는 물체의 거리와 크기를 모두 고려할 수 있도록 퍼지 모델로서 구현하고자 한다.

1. 서론

멀티미디어 관련 기술 개발과 더불어 세계 유수의 연구기관 및 방송사에서는 입체 영상의 구현과 관련된 연구가 한창 진행 중이다. 3D 영상 구현을 위한 다양한 기법들 중에서 양안 시차에 의한 입체감 형성은 가장 기본적이고 빈번히 사용되는 방식중 하나이다[1][2][3]. 이러한 양안 시차를 이용한 3D 영상 구현을 위해서는 대상 물체를 여러 대의 카메라를 이용하여 다양한 각도로 촬영하고 이를 다시 합성하는 기술이 사용된다. 이 때 카메라의 위치 조정에 따라 3D 전용 모니터에서 표현되는 입체 영상의 품질이 결정된다. 그러므로 입체감을 가장 잘 나타내는 카메라의 위치를 찾아내는 것은 3D 영상 구현에 있어 중요한 문제이다. 카메라의 위치 제어에는 자동화를 어렵게 하는 몇 가지 제한사항이 있다. 입체감의 인지는 인간의 감각에 의해 이루어지므로 카메라 위치를 제어하기 위한 일정한 수학적 모델을 제시할 수 없다. 따라서 현재는 영상 전문가의 경험을 근거로 카메라의 위치 조정을 수작업으로 행하고 있다. 이러한 방식으로는 매 순간 위치가 변화하는 대상에 대해 실시간으로 카메라 위치를 제어하는 것이 매우 어렵다. 이에 신경망

을 이용하여 카메라의 위치를 제어할 수 있는 제어 시스템을 구성하려고 한다.

신경망은 주어진 정보를 이용하여 학습함으로써 주어진 시스템의 특징을 추출하는데 매우 우수한 능력을 발휘하고 있다[4][5][6][7]. 그러나 신경망의 학습에 사용되는 데이터가 정확하지 않을 경우 학습이 완료된 신경망의 출력에는 오차가 있을 수밖에 없다. 그런데 카메라 제어를 위한 학습 데이터의 획득은 3D 영상 개발자의 수작업과 주관적 판단에 의해 이루어지므로 획득 데이터에는 오차가 존재할 수 있다. 이러한 오차는 데이터 수집 당시의 주변 환경과 심리적 요인에 큰 영향을 받게 된다. 따라서 보다 정교한 성능의 제어 시스템 구성을 위해서는 학습 데이터를 획득할 때 획득된 데이터를 평가하고 오류를 최소화할 수 있는 중간 과정이 필요하다. 그 중간과정으로 퍼지이론을 적용하였다[8][9][10].

영상 개발자들이 획득한 학습 데이터들의 경향을 분석하여 초기에 선정된 데이터 이외의 추가적인 데이터 후보를 추천한다. 그리고 초기 데이터와 추천된 후보 데이터 중에서 최상의 값을 선택한 후 카메라 위치 제어를 위한 신경망 학습에 사용한다. 이러한 데이터 보정 작업을 거치면 부

정확한 학습 데이터로 인해 발생하는 신경망 시스템의 성능 저하를 방지할 수 있다. 이와 같은 알고리즘은 데이터 획득 과정이 불명확하고 신뢰도를 확신 할 수 없는 시스템을 위해 다양하게 적용될 수 있다. 또한 데이터 경향 분석 과정을 자동화하면 사용자별 개별 학습을 통해 개인의 특성에 맞추어 최적화된 시스템을 구현할 수 있을 것이다.

2. 입체 모니터를 위한 카메라 위치 제어 시스템 개발 환경

본 논문에서는 양안 시차를 이용하여 입체감을 표현하는 '다안식-무안경방식'을 사용한다. 그리고 3D 모니터에 출력할 입체 영상은 우주 공간을 배경으로 하는 게임의 진행 화면으로써 실사 촬영이 아닌 가상공간상의 대상 물체를 각기 다른 각도에서 렌더링하고 이를 합성하여 사용한다.

그림 1에서 보는 바와 같이 각기 다른 위치에 있는 카메라로부터 얻은 데이터는 이미지 합성 프로그램을 통해 입체감이 나타나는 3D 이미지로 변환된다. 그리고 이러한 3D 이미지 데이터는 전기적인 신호로 바뀌어 3D 모니터에 전송된다.

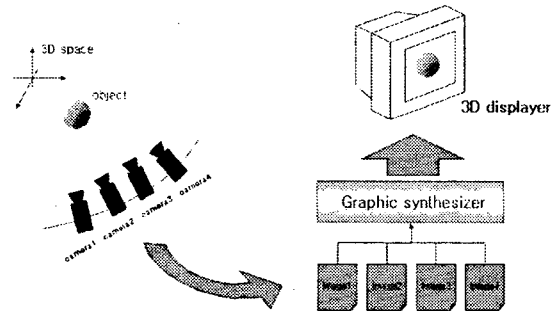
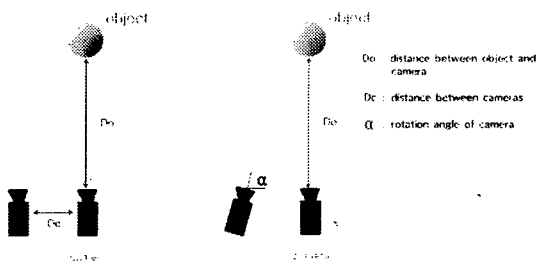


그림 1. 3D 모니터를 위한 이미지 처리

시스템에 따라 카메라의 개수는 다양하게 설정될 수 있으나 본 연구에서는 4대의 카메라에서 각기 4장의 이미지가 추출되는 시스템을 기본으로 하였다.

초점을 조절하는 방법은 두 가지가 있다. 첫 번째는 카메라의 좌우 회전각을 조정하는 것(그림 2. (b))이며 두 번째 방법은 카메라 좌우 회전각은 정면을 주시하도록 고정한 상태에서 카메라 간격을 조정하는 것(그림 2. (a))이다. 여기서 D_0 는 물체와 카메라의 거리, D_c 는 D_0 에 대한 카메라 간격의 비율(카메라간 거리를 d 라 했을 때, $D_c = D_0/d$)이다.



(a) 간격 조정 (b) 회전각 조정
그림 2. 초점 조절 방법

두 가지 방법 모두 물체와 카메라간 거리에 의해 파라미터(회전각 또는 카메라 간격) 값이 결정된다. 그러나 첫 번째 방법은 가상공간을 위해 사용되는 프로그램 특성상 구현이 어려우므로 두 번째 방법을 사용하여 초점을 조정하도록 한다. 신경망 학습을 위한 데이터 획득 과정을 살펴보면 다음과 같다. (그림 3)

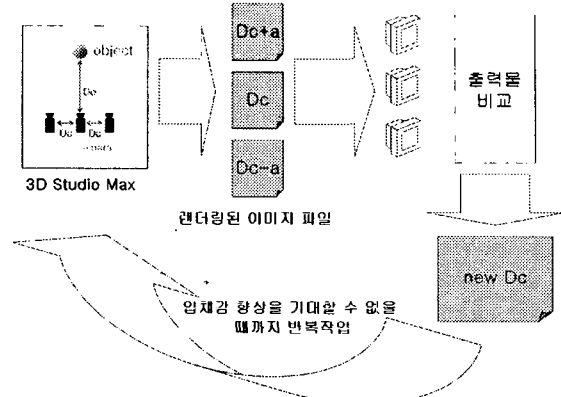


그림 3. 데이터 획득 과정

영상 개발자들은 경험적 지식을 토대로 초기 카메라 간격을 $D_c=20$ 으로 설정하고 입체 영상을 합성한다. 그리고 $D_c=20 \pm a$ (a 는 개발자가 지정하는 임의의 값)일 때의 입체 영상 2개를 추가를 생성한다. 이러한 3개의 입체 영상 중에서 출력 품질이 가장 우수한 영상을 고른 후, 그 영상에서 사용한 D_c 값을 기준으로 다시 추가 이미지를 생성하고 품질을 비교하는 과정을 반복하게 된다. 더 이상 입체감 향상을 기대할 수 없을 때 작업을 종료하고 최종적으로 데이터를 결정하게 된다. 그러나 위와 같은 작업을 반복적으로 수행하게 되면 눈의 피로와 예측하는 심리적인 요인이 작용함에 따라 데이터에 오차가 발생하게 된다. 그러므로 영상 개발자가 발견하지 못한 최상의 카메라 간격 값을 추천할 수 있는 데이터 후보 추천 시스템을 퍼지로서 구현하였다.

3. 학습 데이터 전처리

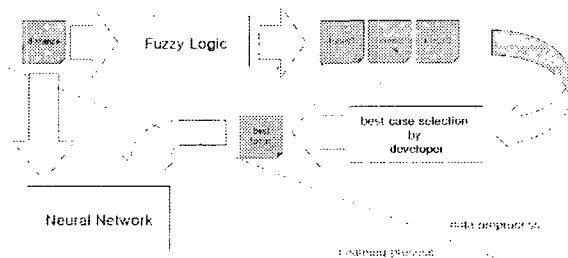


그림 4. 데이터 전처리 및 신경망 학습

카메라 위치 제어를 위한 시스템의 구현은 그림 4와 같이 두 단계의 과정을 거친다. 3D 영상 개발자들은 신경망 학습을 위하여 학습 데이터를 선정한다. 이때 학습 데이터 선정 과정에는 오차가 포함되며 그 원인은 크게 다음 두 가지를 생

각할 수 있다.

- 육안에 의한 주관적 판별
- 학습 데이터 선정 중 심리적 변화

영상 개발자들의 학습 데이터 선정 과정을 살펴 본 결과, 학습 데이터 선정에 있어 신체적인 요인과 심리적인 요인이 매우 강하게 작용함을 알게 됐다. 앞서 찾은 데이터에서의 결과가 현재의 데이터의 결과와 그리 다르지 않을 것이라는 믿음과 눈의 피로도 증가로 인해 더 나은 결과를 찾으려는 노력을 쉽게 포기하는 것이다. 이러한 데이터 선정 오차를 보정하기 위해 일차적인 데이터 획득 후, 중간 처리 과정을 거친다.

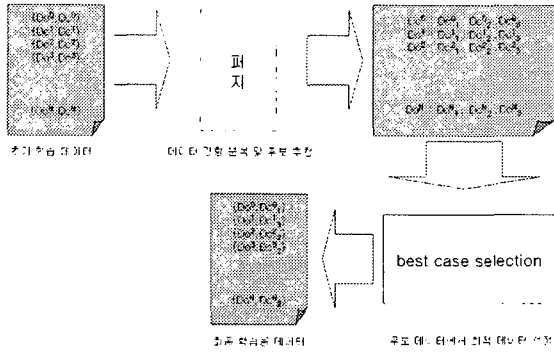


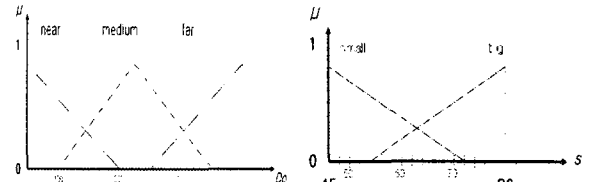
그림 5. 후보 데이터 추천

신경망을 위한 학습 패턴 추출은 전적으로 수작업과 주관적 판단에 의해 이루어진다. 이때, 일차적으로 획득된 학습 데이터는 퍼지를 사용하여 전처리 과정을 거친다. 이는 사용자가 미처 확인하지 못한 데이터에 대한 보완으로써 보다 적절한 파라미터 값을 선택할 수 있도록 도와주는 역할을 하게 하기 위함이다.

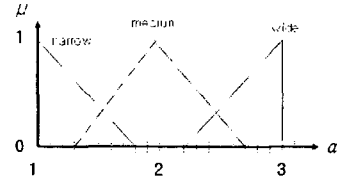
여기서 D_o 는 물체와 카메라의 거리, D_c 는 D_o 에 대한 카메라 간격의 비율(카메라간 거리를 d 라 했을 때, $D_c = D_o/d$)이다. 초기에 획득된 데이터가 위의 과정(그림 5)을 거치면 초기 데이터에서 각각의 학습 쌍에 대해 두개의 후보 초점 거리가 생성된다. 이렇게 중간 처리된 데이터는 영상 개발자들에게 다시 보내지고 세 개의 초점 거리 중에서 최상의 값을 하나씩 선택하여 최종적인 학습 데이터를 선정하게 된다.

입체 영상을 출력할 때 물체의 거리가 가까우면 카메라 간격 조정 범위가 커야만 입체감의 변화를 감지할 수 있다. 그리고 물체의 거리가 먼 경우에는 간격을 조금만 조정해도 입체감의 변화가 뚜렷하게 나타난다. 마찬가지로 대상 물체의 크기가 큰 경우에도 작은 물체일 때보다 상대적으로 카메라 간격 조정 범위가 넓어야 입체감의 변화를 인지할 수 있다. 그러므로 물체의 거리와 크기에 따른 초점 거리 변화를 고려하여 데이터 후보를 추천하여야 한다. 획득된 초기 데이터의 경향을 반영하여 물체의 거리와 물체의 크기에 대한 퍼지 멤버십 함수를 설정하였다(그림 6).

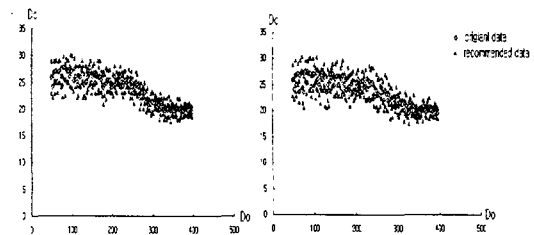
퍼지 전처리에는 Mamdani의 min-max 기법과 비퍼지화를 위해 무게중심법이 사용되었다. 그림 7는 초기 데이터와 퍼지에 의해 생성된 새로운 데이터들을 나타낸 것이다.



(a) 거리 멤버십 함수 (b) 크기 멤버십 함수



(c) 카메라 간격에 대한 멤버십 함수
그림 6. 퍼지 멤버십 함수



(a) 객체 1의 추천 데이터(크기: 45)
(b) 객체 2의 추천 데이터(크기: 80)

그림 7. 후보 추천된 데이터

초기 데이터 (D_o, D_c)

$$\begin{aligned} \text{후보데이터 } D_{c1} &= D_{c0} + \alpha \\ D_{c2} &= D_{c0} - \alpha \end{aligned} \quad (1)$$

여기서, $\alpha = F(D_o, s)$: 물체의 거리와 크기를 입력으로 하는 퍼지 관계 함수

퍼지 전처리 과정을 거친 후 각각의 학습 데이터에 2개씩의 새로운 카메라 간격이 추가되었다. 식 1 이 데이터를 영상 개발자들에게 전달하여 원본과 후보 데이터 사이에서 입체감이 가장 잘 표현되는 카메라 간격을 선택하게 하였다.

표 1. 추천 데이터와 원래 데이터의 사용 비교

	Small object	Big object
총 데이터 개수	117	117
원래 데이터	68	85
추천된 데이터	49	32

전처리 과정을 거친 후, 생성된 데이터의 실제 사용량은 표 1에서 보는 바와 같다. 물체의 크기가 작은 경우에 보다 선택 빈도가 높은 것을 알 수 있다. 이러한 결과가 나온 것은 물체가 작을수록 입체감 판단이 힘들어지기 때문에 학습 데이터 선정을 위한 테스트를 쉽게 포기하는 심리적 요인과 눈의 피로도와 같은 신체적 요인에서 비롯되는 것으로 추측된다.

4. 신경망에 의한 카메라 위치 제어

전처리 과정을 거쳐 선정된 최종적인 학습 데이터는 신경망의 학습에 사용된다. 여기서, 카메라 위치 제어를 위한 신경망의 구조로 다층신경망을 사용한다[11][12].

카메라 제어를 위한 신경망은 입력층, 출력층 그리고 한 개의 은닉층을 갖고 있으며 은닉층은 4개의 노드로 구성되어 있다. 신경망에는 시그모이드 함수와 오차 역전파 알고리즘이 사용되었으며 weight 초기화는 0과 1 사이의 수를 무작위로 추출하여 사용하였다.

표 2. 신경망 학습 파라미터 및 결과

학습 데이터 개수	117쌍
초기 학습률	0.1
초기 weight 설정	0 ~ 1 (random)
RMSE	0.086
학습 완료	23 epoch

신경망을 이용한 카메라 제어기의 입력은 물체의 거리가 사용되며 출력은 카메라 간격이 나오게 된다.

출력 Dc는 3D Studio Max라는 프로그램으로 전달되어 초기 상태에서 설정되었던 뷰 프러스텀에 더하여 3개의 뷰 프러스텀을 생성하고 입체 영상으로 합성한다.(그림 8)

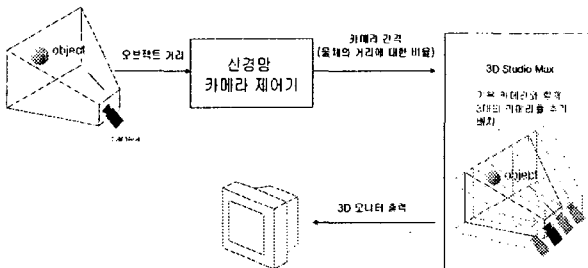


그림 8. 카메라 위치 제어기



(a) 전처리 데이터 결과 (b) 초기 데이터 결과
그림 9. 3D 영상 출력

5. 결론 및 향후 연구

본 논문에서는 신경망 학습을 위한 데이터 획득시 발생하는 오차를 줄이기 위해 학습 데이터를 추천하는 알고리즘을 제안하였다. 3D 모니터를 위한 시스템 설계에 있어 신경망 학습을 위한 데이터 획득 작업은 인간의 주관적 판단에 크게 좌우된다. 그러므로 심리적이고 불확실한 요소로 인해 신경망을 위한 학습 데이터 자체에 오차가

포함되어 제어 시스템 자체의 성능이 저하될 우려가 있다. 이를 극복하기 위하여 퍼지 모델을 사용하여 전문가에 의해 획득된 데이터를 분석하고 새로운 학습 데이터 후보들을 추천하도록 하였다. 그리고 새로 생성된 추천 데이터들과 원래 데이터들 중에서 최상의 데이터를 선택하여 학습을 시킴으로써 수집 과정에서 생기는 오류를 최소화 하도록 하였다.

본 논문에서는 데이터를 획득할 때의 영향 요소를 물체의 위치와 크기로 설정하였다. 그러나 실제에서는 주변의 조명 및 데이터 획득자의 눈의 높이와 자세와 같은 다양한 요소들이 복합적으로 작용한다. 따라서 보다 정확한 학습 데이터 추천 및 획득을 위해서는 위와 같은 요소들의 영향까지 모두 적용시킬 수 있는 보다 심화된 알고리즘의 연구가 필요하다.

감사의 글 : 본 연구는 과학기술부의 뇌신경정보학연구사업에 의해 지원받았습니다.

6. 참고문헌

- [1] <http://www.omnitel.co.kr>
- [2] <http://www.stereo3d.com/3dhome.htm>
- [3] <http://www.3dview.co.kr/sub03/?p=sub01>
- [4] A. Guez, J. Eilbert, and M. Kam, "Neural Network Architecture for Control" IEEE Control Systems Magazine, pp. 22-24, April, 1988
- [5] Kumpati S. Narendra. K. Parthasarathy, "Identification and Control of Dynamic Systems Using Neural Networks" IEEE Trans. Neural Networks. pp 4-27. vol. 1, no. 1, Mar 1990.
- [6] S. R. chi, R. Shouresshi, and M. Tenorio, "Neural networks for system identification" IEEE Contr. syst. Mag, vol. 10, pp. 31-34, 1990.
- [7] J.S.R. Jang, C.T. Sun, E. Mizutani, "Neuro-Fuzzy and Soft Computing", Prentice Hall PTR, 1997.
- [8] George J. Klir and Bo Yuan, "Fuzzy Sets and Fuzzy Logic - Theory and Applications", Prentice Hall PTR, 1995.
- [9] 변중남, Fuzzy Logic Control, 홍릉과학출판사, 1977.
- [10] 이광형, 오길록, "퍼지 이론 및 응용 II", 홍릉과학출판사, 1997.
- [11] Simon Haykin, "Neural Networks - a comprehensive foundation", Prentice Hall PTR 1999.
- [12] Chin-Teng Lin and C.S. George Lee, "Neural Systems - A Neuro Fuzzy Synergism to intelligent Systems", Prentice Hall PTR, 1966.