

한글 지화 인식에 최적화된 데이터 글러브 시스템

민승기[○] 오상혁 김교령 윤태현 임준규 Yunli Lee 정기철
송실대학교 정보과학대학 미디어학과 HCI lab.
dfmin84[○]@nate.com , hyeoki@gmail.com, uki0413@naver.com
{niceyth, liger82, yunli, kcjung}@ssu.ac.kr

Optimize Data Glove-based System for Korean Finger Spelling Recognition

Seungki Min[○], Sanghyeok Oh, Gyoryeong Kim, Taehyun Yoon,
Chungyu Lim, Yunli Lee and Keechul Jung¹

요 약

본 논문에서는 지화 인식에 최적화된 데이터 글러브 기반의 시스템을 제안한다. 제안된 데이터 글러브는 적은 수의 센서로 인식 속도의 향상을 기대할 수 있으며 한글의 지화 인식만을 위한 특수한 목적을 가지고 저렴하게 설계되었다. 그에 따라 한글의 지화를 사용한 많은 어플리케이션에 쉽게 적용할 수 있을 것이 기대된다. 2개의 틸트 센서는 손의 방향을 인식하고 5개의 플렉스 센서는 각 손가락의 구부러진 정도를 측정한다. 제안된 시스템에서는 k-means 알고리즘과 간단한 인덱싱 방식을 사용하여 한글의 기본적인 음소 24개를 인식하는 실험을 하였으며 인식율은 80.27% 에 이르렀다.

1. 서 론

사람과 사람 사이의 의사소통은 언어적 표현과 비언어적 표현 등 여러 가지 요소들로 구성되어 있다. 이러한 여러 요소들 중, 비언어적 표현은 의사 전달력에 있어서 높은 부분을 차지하고 있다. 실제로 의사소통 과정에서 언어적 표현의 이해는 전체의 35%만을 차지하고 나머지 부분은 표정, 손짓, 몸짓과 같은 비언어적 표현이 대부분을 차지하고 있다 [1].

수화는 이러한 비언어적 표현으로 이루어진 하나의 언어이다. 손의 방향, 위치, 동작과 손가락의 모양을 통해 약속된 의미를 가진 언어로써 사람은 수화를 인식하게 된다. 수화는 동적인 손 제스처가 들어간 수화와, 손의 형태만을 인식하는 지화로 나뉘게 된다. 지화는 한글의 자음, 모음을 표현하는 음소단위로 글자를 표현한다. 주로 이름, 상호, 번호 등 고유명사를 표현하는데 많이 사용된다.

수화의 인식을 위한 데이터수집방법으로는 글러브 기반과 영상기반이 있다. 영상을 기반으로 하는 방법은 사용자에게 어떠한 기구적 부담도 가지 않는다는 장점이 있지만 아직 몇 가지 문제점이 존재한다. 가령 하나의 카메라로 입력을 받을 경우 교합이 생겨서 정확한 인식이 힘들고 환경에 의존적인 경우가 많이 생긴다. 그에 비하여 글러브 기반의 수화인식은 정확한 데이터의 입력으로 인하여 영상기반의 인식방법보다 데이터처리가 쉽고 인식율이 높은 편이다.

현재 개발된 데이터 글러브로는 cyber glove, power glove, data glove, space glove 등이 있다[2-3]. 이런 글러브들은 많은 어플리케이션의 입력장치나

인터페이스로서 사용되고 있다. 하지만 이 글러브들은 적어도 15개 이상의 센서를 사용하고 있다. 센서의 개수가 많아지면 많아질수록 처리해야 할 데이터 양도 증가하게 되므로 시스템의 인식속도는 더 느려지게 되고 데이터 글러브 자체의 가격도 올라갈 수 밖에 없을 것이다. 현재 시중의 글러브들은 수화 인식을 위해 개발된 것이 아닌 다른 용도로 사용되고 있다. 제안된 글러브는 한글 지화의 인식만을 위한 글러브로써 최소한의 센서를 사용하여 최적화 시켰다. 모든 지화는 3방향의 손 동작으로 구분이 가능하고 각 손가락에 하나의 플렉스 센서를 사용함으로써 표현이 가능하다. 손의 방향에 따른 그룹을 2개의 틸트 센서로 구분할 수 있고 5개의 플렉스 센서로 손가락의 형태를 인식할 수 있다.

데이터 글러브를 이용한 수화인식의 연구는 활발히 진행되고 있다. Cemil Oz 와 Ming C.Leu[4]는 cyber glove 와 flock of birds 3D모션 트래커를 사용하고 인공 신경망을 통하여 26개의 ASL(American sign language) 알파벳으로부터 96%의 인식율을 얻었다. Honggang Wang [5]은 26개의 알파벳과 36개의 기본적인 손의 형태로부터 은닉 마르코프 모델을 사용하여 95%의 인식율을 얻었다. 이 연구 역시 입력장치로는 cyber glove 와 flock of birds 3D 모션 트래커를 사용하였다. Jose L. Hernandez-Renollar[6]는 Acceleglove 를 사용하여 22개의 수화로부터 99.3%의 인식율을 얻었다.

이에 본 논문에서는 지화 인식에 적합한 글러브 기반의 시스템을 제안한다. 제안된 글러브는 2개의 틸트 센서와 5개의 플렉스 센서로 입력을 받기 때문에 다른 글러브에 비하여 저렴하고 데이터 처리의 부담도 줄어들게 된다. 2개의 틸트 센서는 손의 방향을

측정하여 k-means 알고리즘을 통해 24개의 음소를 3그룹으로 나누고, 5개의 플렉스 센서는 각 손가락의 굽혀진 정도를 측정하여 양자화 인덱싱을 통하여 데이터베이스와 비교하여 인식한다. 그에 더하여 블루투스를 사용함으로써 사용자가 느끼는 입력장치에 대한 부담도 최소한으로 줄어 들 수 있게 하였다. 제안된 글러브는 지화를 이용한 많은 어플리케이션에서 사용할 수 있을 것으로 기대된다.

본 논문의 구성은 다음과 같다. 2장에서는 제안된 글러브의 전체적인 디자인과 각 부분의 기능, 시스템의 구조를 설명하고, 3장에서는 k-mean 알고리즘과 인덱싱을 사용한 시스템에 대하여 알아본다. 4장에서는 제안된 글러브와 시스템의 실험결과와 실험결과에 대한 분석을 논하고, 마지막으로 5장은 이후 연구와 결론을 맺는다.

2. 지화 인식 시스템

본 논문의 지화 인식 시스템은 제안된 글러브를 바탕으로 이루어진다. 전체적인 시스템의 구성도는 다음과 같으며, 그림 1 (a)는 데이터 글러브에서 데이터의 흐름이고, (b)는 입력장치를 통해 받은 신호가 어떻게 인식되는지를 나타내어주고 있다.

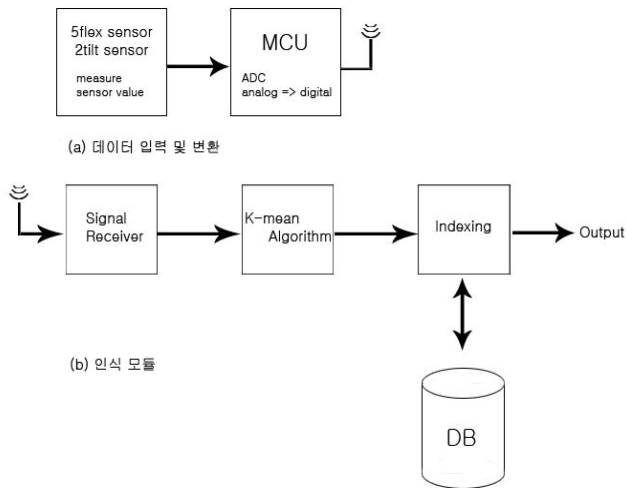


그림 1. 지화 인식 시스템의 구성도

처음에 사용자의 움직임은 센서로부터 측정되어 MCU(micro controller unit)로 보내지게 된다. 이 데이터는 아직 아날로그 신호이지만 MCU를 통하여 디지털 신호로 바뀌게 된다. 이 신호는 블루투스 모듈을 통하여 시스템으로 보내어 지고 시스템은 이를 이용하여 인식을 시작하게 된다.

그림 3 에서 보듯이 지화의 특성상 모든 음소는 방향성이라는 특징을 통하여 4그룹으로 나눌 수 있다. 우리는 먼저 2개의 틸트 센서로 4개의 그룹을

생성하였다. 이 과정에서는 K-means 알고리즘을 사용하였다. 그룹이 분할되면 각 손가락에 설치된 플렉스 센서의 값을 인덱싱 하여 데이터베이스에 있는 값들과 매칭하여 인식을 한다.

2.1 글러브 디자인

제안된 글러브의 목적은 어플리케이션의 입력장치로서 누구나 사용가능하고 부담 없이 접근하는 것에 있다. 다른 글러브에 비하여 저렴한 비용과 적은 수의 센서로 지화 인식 시스템뿐만 아닌 다른 어플리케이션에도 적용하기 쉬울 것이다.

글러브의 재질은 두 가지로 구성되어 있다. 안 쪽은 면 재질의 장갑으로 사용자를 감전으로부터 보호하며 편안함을 느낄 수 있게 하였고 바깥 쪽에는 회로와 센서를 부착시킬 수 있는 합성가죽의 재질로 구성하였다. 각각의 손가락에는 플렉스 센서가 위치하여 손가락의 구부림 정도를 측정하며 손등에 위치한 틸트 센서는 손의 가로와 세로 방향을 인식한다. 블루투스 센서와 MCU는 글러브에 연결된 밴드에 설치되어 착용감에 이상이 없게 하였다.

플렉스 센서는 손가락의 구부림 정도에 따른 저항의 변화를 측정한다. 제안된 글러브의 각 손가락에 설치된 플렉스 센서는 전혀 구부러지지 않았을 때 10K ohm, 90도 구부러졌을 때 3.5K ohm 값을 가진다. 이 값들은 정해진 인덱싱 규칙에 따라 분할되어 LUT(look up table) 에 저장된다.

틸트 센서는 손의 방향을 측정한다. 측정할 수 있는 최대 범위는 -60° 에서 60° 이고, 측정 오차는 0.1° 이다.

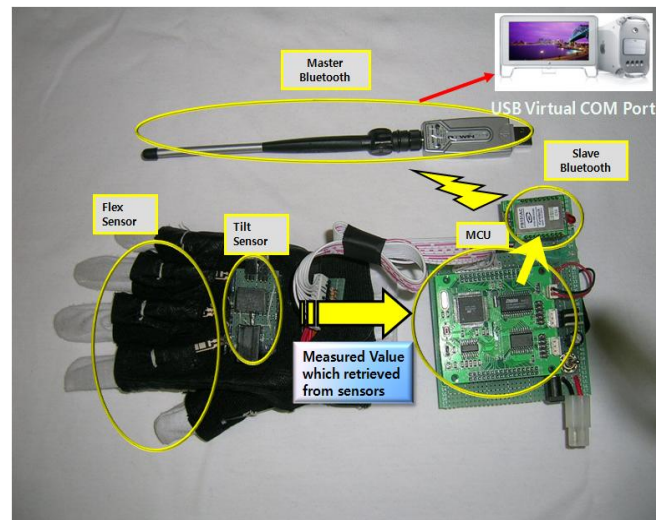


그림 2. 제안된 글러브의 시스템

2.2 데이터입력과 사용자 정보 측정

글러브를 착용한 사용자가 움직이기 시작하면 측정된 데이터는 실시간으로 입력 된다. 이 데이터는 AD Converter를 통해 아날로그 신호에서 디지털 신호로 변환되어 ATmega128을 거쳐 블루투스 모듈로 전송된다. 각 신호는 4byte 이고 7개의 센서로부터 28byte 에 1byte 의 검측 데이터를 더하여 29byte 로 블루투스를 통하여 시스템에 전송된다.

각기 다른 사용자 손의 정보를 일반화 하기 위하여 시스템은 전처리 과정을 필요로 한다. 이 과정을 통하여 얻어진 정보는 데이터베이스에 입력되어 인식 과정에 사용되게 된다.

3. 한글의 지화 인식

한글은 현존하는 언어 중 가장 많은 발음을 표기할 수 있는 언어체계이다. 한글의 기본적인 구성은 14개의 자음과 10개의 모음으로 이루어져 있다. 24개 이상의 음소가 존재하지만 제안된 시스템은 가장 기본적인 24개의 음소만을 대상으로 한다[그림 3].



그림 3. 한글의 기본적인 24 가지 음소

그림 3 에서 보듯이 몇몇 동작들은 식별하기 어려운 형태를 띄고 있다. 3장에서는 이러한 지화들을 어떻게 식별하고 인식하는 지에 대하여 기술한다.

3.1 특징 선택과 추출

제안된 수화 인식 시스템은 각기 다른 사용자 손의 정보를 일반화 시키기 위하여 전처리 과정으로

센싱이라는 과정을 거치게 된다. 센싱은 사용자마다 손의 정보가 다르기 때문에 이를 일반화 하기 위해 구부린 정도를 측정하여 인덱싱 범위를 다르게 하기 위한 과정이다. 10초의 시간 동안 사용자는 각 손가락을 구부리고 피면서 최대 범위를 측정한다. 센싱을 마친 데이터는 실시간으로 계산되어 그 사용자만의 LUT(look up table)에 저장되어 인식에 사용된다.

각 지화를 인식하기 위해 시스템은 29byte 의 데이터를 글러브로부터 입력 받는다. 첫 단계로 우리는 2개의 틸트 센서를 이용하여 24개의 지화를 4개의 그룹으로 분할한다. 손가락을 제외한 지화 동작의 축만으로도 그룹을 나눌 수 있기 때문에 필요 없는 수행을 줄여줄 수 있다. 이에 본 논문은 그룹을 나누는데 빠른 수행을 보이는 K-means 알고리즘을 적용하였다.

그루핑 과정이 끝나면, 각 손가락의 플렉스 센서 값을 비교하여 인식한다. 지화의 모든 동작은 손가락은 완전히 펴진 상태, 반쯤 굽혀진 상태, 완전히 굽혀진 상태로 나눌 수 있다. 플렉스 센서의 값의 범위는 최소 400 에서 최대 800까지로 지화의 특성에 따라 이를 분할하여 3단계로 나누어 사용한다. 이로써 각 지화 동작은 다섯 개의 플렉스 센서 값으로 표현되게 된다. 이 과정을 통하여 우리는 시스템의 속도를 개선하였다.

3.2 K-means 알고리즘의 적용

K-means 알고리즘(KM)은 1967년 J.MacQueen 이 설계하였다[7]. KM 은 모든 데이터를 k개의 그룹으로 분할한다. 각 그룹은 그 그룹의 모든 데이터로부터 계산된 중심점을 갖고 있으며 단계가 적용될 때마다 그룹의 값들과 계산되어 조금씩 변화한다. 몇 번의 단계가 지나고 나면 중심점은 변하지 않고 최적화된 그룹이 생성된다.

그림 3 에서 모든 지화의 동작은 2개의 틸트 센서로부터 3개의 그룹으로 나뉘어 질 수 있다. 'ㄱ', 'ㄴ', 'ㅁ' 이 그 예이다. 이렇게 나뉘어진 동작들은 각 그룹의 고유 값을 가지게 되고 시스템에서는 틸트 값 대신 이 그룹 값을 사용한다. 어떠한 입력 값이 들어올 경우 모든 입력 값으로 모든 동작을 검색하는 대신 그룹 값을 사용하여 비교할 대상을 줄이기 때문에 속도는 비약적으로 증가하게 된다.

3.3 양자화 (quantization)를 통한 인덱싱인덱싱

적은 수의 센서를 사용하지만 값의 범위가 큰 5개의 플렉스 센서 값을 순차적으로 비교하여 인식한다는 것은 시스템에 많은 부담을 주게 된다. 많은 수의 지화 동작들을 인식하기 위해서 우리는 간단한 양자화를 통한 인덱싱 값을 사용하기를 제안하였다.

표 1. 은 본 논문에서 사용한 인덱싱 값을 보여준다. 플렉스 센서 값은 손가락의 구부러진 정도에 따라 3단계의 양자화 과정을 거쳐 3개의 인덱스 값을 가지게

된다. 이 값은 센싱 과정을 통하여 각 사용자의 특성에 따라 정해진 비율로 인덱싱 값을 다시 지정해준다.

표 1. 엄지손가락의 인덱싱 예

센서 값	인덱싱 값
540 ~ 590	0
591 ~ 691	1
692 ~ 741	2

표 2. 는 데이터베이스에 저장되어 있는 지화의 인덱싱 값들이다. 데이터베이스의 인덱싱 값들은 정확한 지화의 동작을 샘플로 만든 정형화된 모델이다. 플렉스 센서의 값 대신 인덱싱 값을 사용함으로써 인식 속도를 한층 더 빠르게 만들 수 있다.

표 2. 데이터베이스의 지화 인덱싱 샘플

손가락 \ 자음	1	2	3	4	5
ㄱ	2	2	0	0	0
ㄴ	2	2	0	0	0
ㄷ	0	2	2	0	0
ㄹ	0	2	2	1	0
ㅁ	0	0	0	0	0
ㅂ	0	2	2	1	2
ㅅ	0	2	2	0	0
ㅇ	1	1	2	1	2
ㅈ	2	2	2	0	0
ㅊ	2	2	2	1	0
ㅋ	2	0	2	0	0
ㅌ	0	2	2	1	0
ㅍ	0	0	0	0	0
ㅎ	2	0	0	0	0

손가락 \ 모음	1	2	3	4	5
ㅏ	0	2	0	0	0
ㅑ	0	2	2	0	0
ㅓ	0	2	0	0	0
ㅕ	0	2	2	0	0
ㅗ	0	2	0	0	0
ㅛ	0	2	2	0	0
ㅜ	0	2	0	0	0
ㅠ	0	2	2	0	0
ㅡ	0	2	0	0	0
ㅣ	0	0	0	0	2

4. 실험결과 및 분석

4장에서는 제안된 지화 인식 시스템의 성능과 인식율을 측정하고 이에 대한 분석을 하였다. 실험은 수화를 접해보지 않은 5명의 남녀가 3번씩 시행하였다.

제안된 글러브와 지화 인식 시스템은 그림 4 와 같다. 제안된 글러브를 착용한 채 사용자는 동작을 취하고 시스템에 연결된 블루투스를 통하여 신호가 전달된다. 각 실험자는 센싱을 통해 개인의 특성을 일반화 하며 실험을 시작한다. 센싱은 실험자로부터 글러브에 익숙해지도록 하는 목적도 포함되어 있다.

그림 5 는 지화 인식 시스템 모듈의 모습이다. 레벨 부분의 “All” 버튼은 센싱의 시작을 의미하는 것이고 엄지손가락부터 10초씩 데이터를 측정하기 시작한다. 이 과정 후 “Apply” 버튼을 누르게 되면 측정된 데이터가 실시간으로 적용되고 데이터베이스에 저장되어 인식에 사용된다. 그림 5 는 ‘ㅋ’ 이 인식된 모듈의 모습이다.



그림 4. 지화 인식 시스템과 제안된 글러브의 실험장면

지화 인식 시스템을 통해 수화를 처음 접해 보는 실험자를 통해 산출된 시스템의 평균 인식률은 80.27% 이다. 이미 입증된 다른 연구에 비하여 낮은 인식률을 갖는다. 이는 그림 6 에서 보듯이 특정 음소의 낮은 인식률로 인하여 발생한다. ‘ㄹ’ 과 ‘ㄷ’ 의 경우 동일한 틸트 값과 인덱싱 값을 가지게 되어 ‘ㄷ’ 의 인식률이 저조하게 산출되었다. 또 ‘ㄹ’ 과 ‘ㅅ’, ‘ㅋ’ 의 경우도 제안된 시스템의 처리 미숙으로 낮은 인식률을 갖게 되었다.

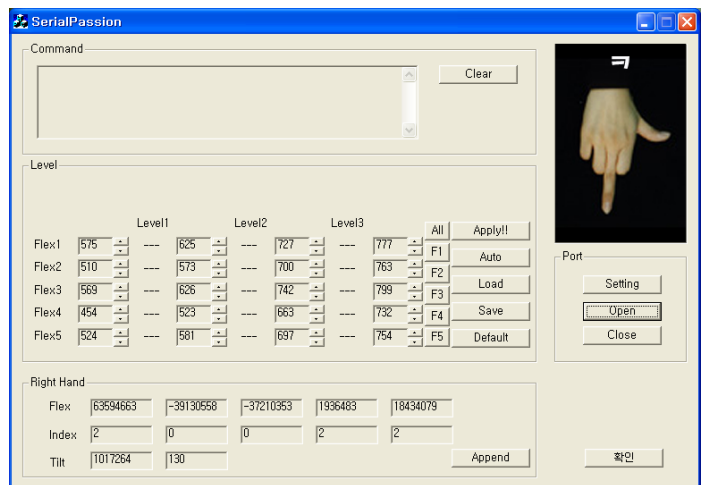


그림 5. 지화 인식 시스템 모듈

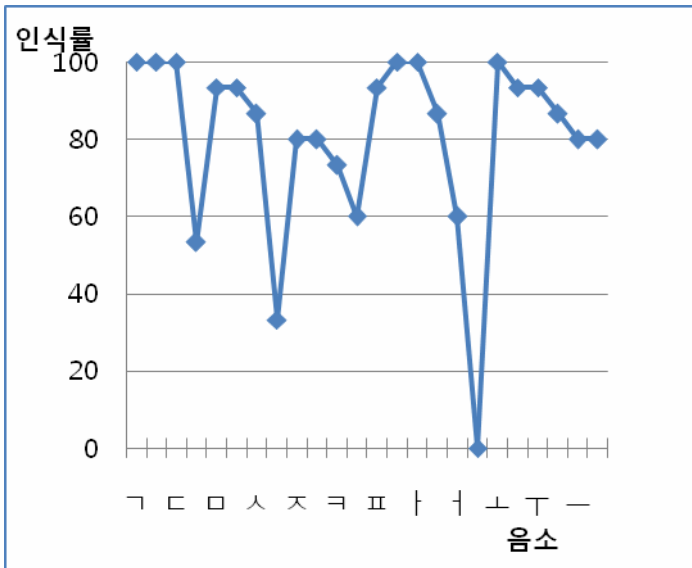


그림 6. 제안된 시스템의 인식 결과

5. 결론

본 논문에서는 직접 제작한 글러브를 통한 한글의 지화 인식 시스템을 제안하였다. 제안된 글러브는 지화 인식 시스템에 필요한 최소한의 센서로 제작 되었으며 이를 통해 인식 속도를 개선하고 가격을 내려 누구나 쉽게 접하고 사용할 수 있게 하였다. 시스템 자체의 성능도 적은 데이터에 유용한 알고리즘을 사용함으로써 한층 더 빠르게 하였다. 이런 간단한 글러브는 다른 어플리케이션에서도 어떠한 전처리 없이 간단히 사용할 수 있기 때문에 현재 존재하는 어떤 글러브보다 확장이 용이하다. 특히 한글의 수화 인식 시스템의 경우 현재 최적화 되어 있기 때문에 사용이 용이할 것이라고 보여진다.

그러나 현재 제안된 글러브는 몇몇 지화에 대하여 올바른 인식을 하기 힘들게 제작되어 있다. 부착된 7개의 센서의 값이 같을 경우, 예를 들어 ‘ㄹ’ 과 ‘ㄷ’ 의 경우가 그 예이다.

이후 우리는 이 문제를 해결하기 위하여 압력 센서를 손가락 사이에 부착하여 이 문제를 해결하고자 한다. 그리고 초음파 센서를 부착시켜 틸트 센서로 구분할 수 없는 손의 방향을 식별할 것이다.

참고 문헌

[1] 송효섭, 장세진, 신병주, 양윤모, “손의 형상과 움직임 방향 정보를 이용한 수화 인식” 한국정보과학회, 6.1996.

[2] Kostas N. Tarchanidis and John N. Lygouras. “Data Glove With a Force Sensor” IEEE

Transactions on Instrumentation and Measurement, VOL. 52, NO. 3. 2003.

[3] David J.Sturman and David Zeltzer. “A Survey of Glove-based Input. Computer Graphics and Applications” IEEE, 1994.

[4] Cemil Oz and Ming C. Leu. “Recognition of Finger Spelling of American Sign Language with Artificial Neural Network Using Position/Orientation Sensors and Data Glove” ISNN, 2005.

[5] Honggang Wang, Ming C. Leu and Cemil Oz. “American Sign Language Recognition Using Multi-dimensional Hidden Markov Models” Journal of Information Science and Engineering 22, 2006.

[6] Jose L. Hernandez-Rebollar, RobertW. Lindeman, Nicholas Kyriakopoulos. “A Multi-Class Pattern Recognition System for Practical Finger Spelling Translation” Fourth IEEE International Conference on Multimodal Interfaces, 2002.

[7] Vitorino Ramos and Fernando Muge. “Map Segmentation by Colour Cube Genetic K-Mean Clustering” Proceedings of the ECDL. 2000.