

TVML을 기반으로 하는 살아있는 실감 콘텐츠의 구현

이관욱 길종인 김제동 김만배*
강원대학교 IT대학 컴퓨터정보통신공학과
{lkw1340, jigit, jedongkim, manbae*}@kangwon.ac.kr

Producing Alive Realistic Contents from a Single Painting based on TVML

Kwan Wook Lee, Jong In Gil, Jedong Kim and Manbae Kim
Dept. of Computer and Communications Engineering, IT College, Kangwon National University

요약

실감콘텐츠는 3DTV, 다시점 TV 등의 필수적인 콘텐츠로, 향후 다양한 유형의 콘텐츠 제작이 필요하다. 본 논문에서는 3D 환경의 방송용 콘텐츠로 살아있는 콘텐츠의 제작기법을 제안한다. 살아있는 콘텐츠는 등장하는 캐릭터가 이동하고, 얘기를 하도록 제작되어, 기존의 단순한 2D영상의 시청에서 살아있는 캐릭터를 시청할 수 있다. 제안 시스템은 TVML (TV program Making Language)을 기반으로 제작되며, 한 장의 영상이 주어지면, 3D Studio MAX 등과 같은 3D 콘텐츠 제작 툴을 이용하여 캐릭터 및 배경, 물체에 대한 객체 화일을 생성한다. 이렇게 제작된 객체 화일들을 TVML 스크립트에서 불러오고, 이들의 동작을 지시함으로써 살아있는 3D 영상 콘텐츠를 제작한다. 실험 콘텐츠로 신윤복의 단오풍정 그림을 적용하여, 배경에 깊이를 주고 또한 캐릭터들이 움직이고 말하게 된다. 사용자가 다양한 이벤트를 제작하면 보다 실감 있고 관람객의 몰입도를 높일 수 있다. 개인이 제작하기 힘든 방송용 실험 콘텐츠를 손쉽게 제작함으로써 Web 2.0 환경에서의 사용자 참여 중심의 콘텐츠 제작의 용이성을 제공한다.

1. 서론

최근 3D display의 지속적인 개발과 공급으로 최근 이를 활성화 할 수 있기 위해 실감영상 제작의 중요성이 증대되고 있다. 실감영상은 2대의 입체카메라를 이용하여 입체영상으로 획득할 수 있고, 특수효과 영화에서 자주 사용되는 그래픽기술을 이용하여 3D 그래픽영상을 제작하기도 한다. CG와는 달리 실사영상을 획득해주는 입체카메라의 사용은 상당히 복잡한 작업을 필요하기 때문에, 장시간의 영화나 콘텐츠 제작은 시간과 인력을 필요로 하고 있다. 스테레오 카메라를 이용하는 입체 콘텐츠의 어려움을 극복하고, 또한 기제작된 2-D 콘텐츠를 입체화하는 입체변환 기법이 영화사를 중심으로 활발히 진행되고 있다. 또한 다시점 모니터의 공급으로 다시점 콘텐츠의 제작도 관심을 받고 있다. 상기 다양한 입체 콘텐츠 제작에도 불구하고, 실감방송의 발전을 위해 새로운 실감콘텐츠의 개발이 필요한 실정이다.

이러한 추세에 맞추어, 기존 2D TIP [2, 3, 4]와 3차원 입체 Tour-into-Picture가 제안되었다 [4]. 실제 TIP는 지속적으로 관심이 대두되고 있다. 상기 연구는 2-D 영상에 국한되어 있는 기존의 TIP를 3차원 입체로 구현한 3D TIP이다. 이러한 입체 TIP는 기존의 2D영상으로만 시청하던 콘텐츠를 3D 입체로 시청하여 보다 몰입감과 실감적인 콘텐츠를 제공하는 장점이 있다. 또한 최근에 Education, IT Technology, Entertainment가 결합된 Edutainment 형 테마 전시관 작품 속 인물과 공간을 실제로 재현하여 작품을 직접 보고 대화할 수 있는 체험형 전시 공간 구성 기법이 제안되었고 [5], 이는 교육부 제 7차 교육과정에서 채택된 펠드먼의 4단계 감상법을 도입한 국내 초유의 감상교육 기술이다. 이렇게 실감콘텐츠는 지속적으로 새로운 형태가 연구되고 있는데, 이러한 추세에 맞추어 본 논문에서는 3D TIP를 살아있는 (alive) 미술관 기법을 적용하여 새로운 실감 콘텐츠 제작 기법을 제안한다. 이를 위해서 TVML (TeleVision program Making Language) [6, 7]을 이용하여 2D 콘텐츠에 살아 있는 캐릭터를 만들고, 오디오를 활용한다. 실험

영상으로 한국 고전화인 신윤복의 '단오풍정'을 선택하였다. 단오풍정에는 다양한 인물들과 나무 및 그네들이 있어 구현에 적합한 영상이다.

본 논문의 구성은 다음과 같다. 2절에서는 관련 연구를 소개하고, TVML 구조 및 기능을 소개하고, 이어서 3절에서는 제안하는 '살아있는 단오풍정'을 제안한다. 마지막으로 4절에서는 결론을 맺는다.

2. 관련 연구

가. 3D TIP

그림 1은 전처리, 렌더링, 후처리 및 입체영상 생성의 4 단계로 구성되어 있는 3D TIP의 전체 흐름도를 보여준다 [4]. 전처리에서는 입력영상으로부터 전경 마스크, 배경영상, 및 깊이맵을 생성한다. 렌더링에서는 이동과 회전이 가능한 가상 카메라의 투영된 텍스처 영상 및 깊이맵을 생성한다. 후처리 과정에서는 전경객체들을 입체로 시청할 때 발생하는 카드보드 효과를 감소하기 위해, 전경객체 깊이에 불륨을 주기 위한 필터링 기법을 사용한다. 마지막으로 입체영상 생성에서는 전경객체와 배경영상의 깊이맵을 이용하여 입체영상을 생성한다.

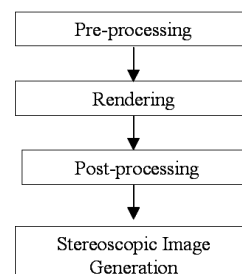


그림 1. 3D TIP의 전체 흐름도

나. 살아있는 미술관

살아있는 미술관[5]은 ‘새로운 미술 감상 교육 전시’, ‘어린이부터 성인까지 쉽게 이해하고 감상할 수 있는 교육전시’, ‘세계최초, 서양미술사가 테마로 구성된 전시’라는 슬로건을 앞세워 예술적, 역사적으로 가치가 중요시 되는 61 점을 선별하여 5개의 흥미로운 테마로 구성하였으며 또한 고대부터 현대까지 그림 속에서 시간 여행을 즐길 수 있다.

총 5개의 테마로 제작되어 있는데, 첫 번째로는 고대 미술 속에 담겨져 있는 수술적 메시지와, 평면적 표현의 이해를 돕기 위해 독특한 퍼포먼스가 펼쳐지는 Story Line을 구성 하여 2D애니메이션과 함께 연출한다. 두 번째로는 대형 특수홀로그램 (와이드매직비전 기술 적용)을 이용한 리얼한 3D Animation, 홀로그램 영상으로 조각상을 실감 나게 연출, 살아 움직이게 한다. 세 번째는 평면적이고 장식적인 그림이 아닌 3차원의 입체감과 공간감으로 표현된 사실적인 그림들을 살펴볼 수 있다. 또한 원근법이 본격적으로 그림에 나타나기 시작한 시기의 대표작들을 디오라마 + 와이드 매직비전 영상으로 연출 제작하고, 그림 속 인물들의 배치와 배경에 대한 작가의 설명을 연출한다. 여기에는 액자 형식의 모니터로 동영상용, 액자 속의 모니터자가 관객들과 직접 대화를 나누는 실감적인 콘텐츠도 포함된다. 네 번째, 명화를 단편애니메이션 화 된 영상으로 재현하여 정지된 명화의 장면 그 이전, 이후의 궁금함을 역사적 고증을 통한 3~5분 풀 애니메이션으로 제작하여, 그 시대의 상황들을 다큐멘터리 형식으로 연출한다. 이는 작품 감상의 극대화를 위해 오감을 동원한 연출기법으로 재구성, 보다 실감나는 작품을 감상할 수 있다. 끝으로 다섯번째에는 정형화된 형태가 작품에서 어떻게 재구성되어지는지를 특수모니터로 재현하게 된다.

개발되어진 콘텐츠들은 Direct X9를 기반으로 한 Real Time Digital Puppet 환경에서 개발되었으며, 동작 및 애니메이션은 립싱크 구동 프로그램 등을 이용하여 개발하였다. 또한 캐릭터는 노멀맵, 카툰 렌더링, 웨도우 맵 등을 사용하여 사실적으로 구현하였으며 사용한 소프트웨어로는 Lipsink SDK(di-o-matic) / 자체 개발 실시간 구동 엔진을 이용하여 개발하였다. 이를 통하여 액자 속의 모니터자가 정교한 모션 캡처로 관객들과 직접 실시간 대화를 나누는 등 단순히 일차원적으로 관람하는 것에서 한발 더 나아가 관람객과 대화하며 관람하는 새로운 문화를 창출해 내었다. 이는 단순한 교육에서 벗어나 관람객의 몰입도를 높이고, 쉽게 접하고 경험함으로써 기억에 오래 남게 되며 보다 효과적인 교육 방식의 채택으로 딱딱하게 느껴지는 미술 감상에 있어서 쉽고 재미있게 유도할 수 있다는 점에서 매우 흥미롭다.

다. TVML의 구조와 기능

TVML이란 문자를 동영상으로 변환하기 위한 일종의 스크립트 언어로, 1996년 일본의 NHK에 의해 개발되었으며 컴퓨터 애니메이션과 실시간 3D CG와 음성을 합성하여 손쉽게 방송용 3D 콘텐츠를 제작할 수 있는 언어이다 [6, 7, 8].

최근의 미디어 산업은 빠른 속도로 성장하고 있으며, 인터넷 기반의 커뮤니케이션, TV 방송 등 여러 분야가 서로 합쳐져 새로운 분야의 산업이 생성되고 있다. 이러한 미디어가 성숙함에 따라 일반 인터넷 유저들도 콘텐츠를 생성하고 또 이를 배포할 수 있게 된다. 최근 각광받고 있는 YouTube, UGC(혹은 UCC) 등이 이러한 추세에 따라 새로이 생성된 문화라 할 수 있다.

그러나 아직 일반인이 제작하기에 콘텐츠 제작 방법은 많이 부족하

며 어려운 것이 사실인데, 일반인은 프로그래밍에 대한 지식이 부족할 뿐 아니라 촬영 장비, 자금, 장소 등 많은 제약조건이 따르기 때문이다. 이러한 어려움에 따라 일반인도 손쉽게 3D 방송용 콘텐츠를 생성할 수 있도록 개발 된 것이 바로 TVML이다. TVML은 일종의 스크립트 언어로 이루어져 있는데, 사용자가 대사를 쓰듯이 TVML 언어로 작성한 후 NHK에서 만든 플레이어를 이용하여 작성된 스크립트를 읽고 분석하여 즉각적으로 실시간 CG나 음성 합성 기술을 이용하여 3D 영상을 생성하게 된다.

이 TVML의 세 가지 기능은 바로 UGC, Automatic Video Production, 그리고 Interactive Contents이다. 첫째, UGC(User Generate Contents) 기능으로, 이는 위에서 설명한 것처럼 일반인이 손쉽게 콘텐츠를 제작할 수 있는 것이다. 둘째, Automatic Video Production이다. 이는 TVML 언어를 이용한 스크립트를 제작하게 되면 이를 분석하여 자동으로 3D 콘텐츠를 제작할 수 있는 것이다. 셋째, Interactive Contents는 바로 살아있는 미술관과 유사한 기능으로, 제작된 콘텐츠는 관람자를 중심으로 유연하게 스크립트를 변경하여 실시간으로 3D 콘텐츠를 제작할 수 있게 되며, 관람자와 대화 하는듯한 기능을 줄 수 있다.

TVML은 크게 선언과 조작, 시나리오로 나눌 수 있다. 선언부에서는 3D 툴을 이용하여 모델링 한 캐릭터와 물체, 배경, 음원 등을 불러와 화면에 세팅하여 사용할 준비를 할 수 있으며 광원, 카메라 및 타이틀을 설정할 수 있다. 조작 부분에서는 캐릭터 및 물체, 광원 및 카메라 등의 동작이나 위치, 특징 들을 조작 할 수 있으며 선언부에서 미리 설정해 놓은 값을 불러와 오브젝트들을 조작할 수 있게 된다. 마지막으로 시나리오 부분에서는 캐릭터의 대사 및 움직임, 장면 전환, 카메라 스위칭 등 콘텐츠의 흐름을 설계함으로써 실제적으로 콘텐츠를 제작할 수 있게 된다.

3. 제안하는 살아있는 단오풍정 기법

TVML script 는 TVML player의 입력 데이터이다. 위에서 언급했듯이, TVML 언어는 가상 3D 캐릭터의 특성과 뷰포인트의 제어를 할 수 있다. TV 프로그램과 같은 가상 3D 콘텐츠는 TVML을 이용하여 제작할 수 있다. TVML은 XML과 유사한 형태의 스크립트 언어이기 때문에, 사용자가 3D 캐릭터 콘텐츠를 더욱 쉽게 제작할 수 있도록 한다. Foreground 캐릭터 모델링은 각 foreground 캐릭터를 3D 모델로 만드는 것이다. 단오풍정의 그림에서 10명의 캐릭터에 대해서, 그림 2에서 볼 수 있는 것과 같이 캐릭터 모델에 대한 Object 데이터 파일을 생성한다. 이렇게 생성된 Object 데이터 파일을 이용하여 캐릭터 모델의 동작을 제어할 수 있게 된다.

```
mllib boxx.mtl
# object Box01
v -29.8893 0.0000 30.6273
v -29.8893 0.0000 -29.8893
v 29.8893 0.0000 -29.8893
v 29.8893 0.0000 30.6273
v -29.8893 53.1365 30.6273
v 29.8893 53.1365 30.6273
v 29.8893 53.1365 -29.8893
v -29.8893 53.1365 -29.8893
# 8 vertices

vn 0.0000 -1.0000 -0.0000
vn 0.0000 1.0000 -0.0000
vn 0.0000 0.0000 1.0000
vn 1.0000 0.0000 -0.0000
vn 0.0000 0.0000 -1.0000
```

```

vn -1.0000 0.0000 -0.0000
# 6 vertex normals

vt 1.0000 0.0000 0.0000
vt 1.0000 1.0000 0.0000
vt 0.0000 1.0000 0.0000
vt 0.0000 0.0000 0.0000
# 4 texture coords

g Box01
usemtl wire_255255255
s 2
f 1/1/1 2/2/1 3/3/1 4/4/1
s 4
f 5/4/2 6/1/2 7/2/2 8/3/2
s 8
f 1/4/3 4/1/3 6/2/3 5/3/3
s 16
f 4/4/4 3/1/4 7/2/4 6/3/4
s 32
f 3/4/5 2/1/5 8/2/5 7/3/5
s 64
f 2/4/6 1/1/6 5/2/6 8/3/6
# 6 polygons

```

다음은 3D로 모델링한 캐릭터를 TVML 스크립트를 이용하여 정의한 것이다.

```

// character define :Mina3
character: casting(name=Mina3)
character: openmodel(modelname=c3, filename="mina\mina.bm")
character: bindmodel(name=Mina3, modelname=c3)
character: setvoice(name=Mina3, voicetype="Male Whisper")
character: blink(name=Mina3, switch=on)
character: visible(name=Mina3, switch=on)
character: position( name=Mina3, x=1.9, y=1.6, z=-1.5, d=-50.0,
posture=standing )
character: swing(name=Mina3, switch=on, degree=2.5, speed=0.5)

```

현재 버전의 TVML은 한국어를 지원하지 않기 때문에, 각 캐릭터의 대화 내용을 Wave 파일로 저장해야 한다. Wave 파일로 저장된 캐릭터의 Audio 파일은 TVML 스크립트 상에, TVML 라이브러리가 제공하는 명령을 이용하여 호출하게 된다. 이벤트 발생은 다양한 동작 명령에 의해 만들어진다. TVML은 다음과 같은 캐릭터의 동작에 대한 명령을 지원한다: *position, talk, gesture, walk, step, stop, sit, stand, turn, bow, look, byebye, action, swing*. TVML 스크립트는 TVML player의 입력 데이터이다. 시나리오에 대한 제작은 사용자의 능력이다. 다음의 OBJ script는 모델 데이터의 동작을 기술한 한 예이다.

TVML player를 이용하여 다음의 그림 2와 같이 애니메이션이 제작되어지고, 초기 렌더링 된 3D 영상을 보여주며 8명의 여인 캐릭터와 2명의 동자승 캐릭터, 그리고 그네를 모델링 하였다. 이 중 한 캐릭터와 그네를 렌더링하기 위한 스크립트는 아래에서 보여준다.



그림 2. 초기 렌더링 된 3D 영상

```

set: change(name=NULL)
// Background

```

```

prop: assign(name=prop50)
prop: openmodel(name=prop50, filename="./BG.obj")
prop: visible(name=prop50, switch=off)
prop: position(name=prop50, x=-4.3, y=-1.8, z=-6.8, pitch=0.00, yaw=0.00,
roll=0.00, scale=0.015)

// swing
prop: assign(name=swing)
prop: openmodel(name=swing, filename="./swing.obj")
prop: visible(name=swing, switch=off)
prop: position(name=swing, x=0.5, y=2.0, z=-1.5, pitch=290.00, yaw=0.00,
roll=-60.00, scale=0.006)

// character define :Mina
character: casting(name=Mina)
character: openmodel(modelname=c1, filename="mina\mina.bm")
character: bindmodel(name=Mina, modelname=c1)
character: setvoice(name=Mina, voicetype="Female Whisper")
character: visible(name=Mina, switch=off)
character: position( name=Mina, x=0.2, y=1.15, z=-1.5, d=-50.0,
posture=sitting )

character: look( name=Mina, what=camera, speed=2.36 )

```

그림 3은 3D로 모델링 된 캐릭터의 동작을 연속적으로 보여주고 있다. 이 장면에서는 오른 손을 흔들 수 있도록 'byebye' 명령어를 사용하였고 대사도 입을 움직여 가능하도록 구현하였다. 구현 스크립트는 아래에서 보여준다.



그림 3. 3D로 모델링 된 캐릭터의 동작

```

character: talk( name=Mina, text="Let's tour into picture." )
wait( time=1 )
character: talkfile(name=Mina,filename="BOOMAR1.wav",
lipsensitivity=2.00, caption="Konnichiwa-" )
character: byebye(name=Mina,times=5,speed=2.00,degree=37.50,wait=yes)

```

그림 4는 배경 뒤로 숨은 동자승의 3D 캐릭터를 모델링 한 그림이다.



그림 4. 배경 뒤로 숨은 동자승의 3D 캐릭터 렌더링



그림 5. 배경 위에서 걷기 명령을 수행하는 캐릭터의 연속적인 동작

다음은 걷기 명령을 수행하는 캐릭터의 동작을 지시하는 스크립트이다.

```
character:walk(name=Mina4,x=2.0,z=-2.0,pitch=1.0,compass=21.0,step=1.0,wait=no)
```

TVML에서 표현될 객체는 배경을 바탕으로 하여 Object를 배치하는 방법으로 표현된다. 배경을 표현하기 위해서는 3DMAX에서 디자인하게 되는데, 이는 기존에 [1]에서 생성된 깊이맵을 토대로 하여 디자인한다. 하나의 평면 위에 인물 Object들을 제거한 배경이미지를 Texture mapping시킨 후에, 폴리곤의 정점 좌표들을 생성한 후, 해당하는 좌표값과 깊이값을 할당시킨다. 렌더링한 결과를 그림 6에서 보여준다. 이를 플레이어에서 구동하기 위한 TVML script는 아래에서 보여준다.



그림 6. 배경에 깊이값을 할당한 후의 결과

```
prop: assign(name=prop50)
prop: openmodel(name=prop50, filename="..BG.obj")
prop: visible(name=prop50, switch=off)

skipscript( switch=off )
video:switcher(source=studio)

camera: assign(name=Acam)

prop: visible(name=prop50, switch=on)
super: on(type=text, text="DanoPungJung Background",
color=ff0000, border=on, bordercolor=0000ff)

sound: open(name=music, filename=danopungjung.mp3)
sound: play(name=music)
camera: movement (name=Acam, x=300, y=200, z=700, vangle=60);
wait(time=5)
camera: movement (name=Acam, x=100, y=100, z=500, vangle=50);

prop: transform(name=prop50, tx=-400.0, ttime=5, wait=no)
wait(time=5)
```

이를 TVML에서 출력하기 위해서는 JPEG 배경영상, 3DMAX에서 제작된 파일(OBJ, MTL)을 필요로 한다. 즉, TVML에서는 배경영상에 OBJ파일의 정보를 할당하여 3D로 표현하게 되는데, camera command를 이용하여 출력할 시점을 지정해주고, 시점변경을 통해 배경을 tracking 할 수 있다. 마지막으로 다양한 카메라 시점과 위치에 따른 영상들은 그림 7에서 보여준다.

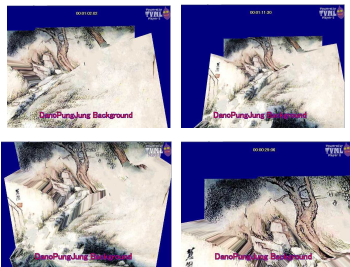


그림 7. 다양한 카메라의 변화에 따른 3D 배경 이미지의 Translation

```
camera: movement (name=Acam, x=300, y=200, z=700, vangle=60);
camera: movement (name=Acam, x=100, y=100, z=500, vangle=50);
prop: transform(name=prop50, tx=-400.0, ttime=5, wait=no)
prop: transform(name=prop50, ty=-300.0, ttime=5, wait=no)
```

```
prop: transform(name=prop50, tz=100.0, ttime=5, wait=no)
camera: movement (name=Acam, x=350, y=0, z=700, vangle=50);
prop: transform(name=prop50, rx=50.0, rtime=5, wait=no)
prop: transform(name=prop50, ry=50.0, rtime=5, wait=no)
prop: transform(name=prop50, rz=50.0, rtime=5, wait=no)
```

4. 결론

본 논문에서 제안한, TVML 스크립트 언어를 이용한 3D 콘텐츠 제작에 대한 방법은 비전문가인 일반인도 전문적인 지식의 도움 없이 방송용 3D 콘텐츠를 자동으로 제작할 수 있다는 데에 그 의의가 있다. 또한 연구 과정을 통해 다양한 콘텐츠를 적용시켜 쉽고 자유로운 3D 콘텐츠 제작이 가능함을 확인할 수 있었다.

이러한 TVML의 기술을 응용하여 수많은 한국화의 교육용 콘텐츠 개발로 인한 관심도 증가 및 교육 효과 향상이 기대되며, 향후 외부로부터의 입력을 감지하여 관람자에게 Feedback을 줌으로써 보다 더 실감 있고 Interactive한 콘텐츠 개발을 목표로 하고 있으며, 나아가 다양한 한국화를 이용한 3D 교육용 콘텐츠 개발을 연구 방향으로 삼을 계획이다.

감사의 글

본 연구는 교육과학기술부와 한국산업기술재단의 지역혁신인력양성사업과 지식경제부 및 정보통신산업진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음. (NIPA-2009-(C1090-0902-0017))

참고 문헌

- [1] Y. Horry, K. Anjyo, and K. Arai, "Tour Into the Picture: Using a Spidery Mesh Interface to Make Animation from a Single Image," Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques, pp. 225-232, 1997.
- [2] K. Boulanger, K. Bouatouch, and S. Pattanaik, "ATIP: A Tool for 3D Navigation inside a Single Image with Automatic Camera Calibration," EG UK Theory and Practice of Computer Graphics 2006.
- [3] N. S. Chu and C. Tai, "Animating Chinese landscape paintings and panorama using multi-perspective modeling," Proceedings of International Conference on Computer Graphics, IEEE, pp. 107-113, 2001.
- [4] 조철용, 김만배, "3DTIP: 한국 고전화의 3차원 입체 Tour-Into-Picture," 방송공학회논문지 제14권 제5호, pp.616-624, 2009.
- [5] 강지연, "Alive gallery", HCI학술대회, 2009.
- [6] Hayashi, M., Ueda, H. and Kurihara, T., "TVML (TV program Making Language) - Automatic TV Program Generation from Text-based Script - ", Proceedings of Imagina'99, (1999)
- [7] Hayashi, M., Gakumazawa, Y., Yamanouchi, Y., "Automatic Generation of Talk Show from Dialog using TVML", ICCG'99 Digital Convergence for Creative Divergence Vol. 2, pp. 325-332, 1999.
- [8] TVML Homepage, <http://www.nhk.or.jp/str/tvml/index.html>