

옷감 시뮬레이션의 절차적 사운드 합성을 위한 새로운 사운드의

에너지와 반전 매핑

김동희[○], 문성혁^{*}, 신영찬^{*}, 김종현^{*}

[○]강남대학교 소프트웨어응용학부,

^{*}강남대학교 소프트웨어응용학부

e-mail: jonghyunkim@kangnam.ac.kr

Novel Sound Energy and Reversal Mapping for Procedural Sound Synthesis in Cloth Simulation

Dong-Hui Kim[○], Seong-Hyeok Moon^{*}, Young-Chan Shin^{*}, Jong-Hyun Kim^{*}

[○]School of Software Application, Kangnam University,

^{*}School of Software Application, Kangnam University

● 요약 ●

본 논문에서는 물리 기반 옷감 시뮬레이션에 적합한 소리를 효율적으로 생성하기 위한 데이터 기반 합성 기법을 제안한다. 시뮬레이션에서 소리를 표현하는 방법은 크게 생성과 합성이 있지만, 합성은 실시간 애플리케이션에서 활용이 가능하기 때문에 인터랙티브한 환경에서 자주 활용되고 있다. 하지만, 데이터에 의존하기 때문에 원하는 장면에 부합하는 사운드를 합성하기는 어려우며, 기존 방법은 한 방향으로만 사운드 데이터를 검색하기 때문에 불연속적으로 인한 끊김 현상이 발생한다. 본 논문에서는 양방향 사운드 합성 기법을 제시하며, 이를 통해 불연속적으로 합성되는 사운드 결과를 효율적으로 개선될 수 있음을 보여준다.

키워드: 소리 합성(Sound synthesis), 옷감 시뮬레이션(Cloth simulation), 가상 소리(Virtual sound), 물리 기반 시뮬레이션(Physically-based simulation)

I. Introduction

최근 인터넷의 새로운 성장과 메타버스의 등장으로 인해 가상환경의 발전가능성이 돋보이기 시작했다. 이로 인해 가상환경 내부의 몰입감을 높이기 위한 연구가 여러 분야에서 진행되고 있으며 관심이 높아지는 추세이다[1]. 시간에 따른 청각의 상호작용에 대해 다양한 연구가 진행되는 반면에 실제 사회적용 사례는 거의 찾을 수 없다. 인간의 오감 중 시각정보 다음으로 청각정보를 많이 받아들이기 때문에 영상처리 분야에서 몰입도 증진을 위해 청각 요소는 필수적인 요소이다. 따라서 본 논문에서는 소리 합성을 통한 고품질 소리 생성 기술을 제안한다.

II. Preliminaries

1. Related works

1.1 Physical based sound generation

컴퓨터 애니메이션과 비디오 게임을 위한 절차적 오디오 생성의 중요성은 오랫동안 인식되어 왔다. 대표적인 연구로 강체 시뮬레이션

을 위한 소리의 생성이 있다[2,3]. 금속, 유리 판 등 강체의 경우 대부분의 물질이 균질하고 등방성이기 때문에 외부 충격 하에서 강한 울림 공명을 나타낸다. 따라서 이들의 소리는 선형 물리학 기반 접근법을 사용하여 비교적 간단하게 소리를 모델링했지만[2], 상대적으로 시뮬레이션의 계산 시간이 오래 걸리며 재질 표현이 어렵다.

1.2 Sounds synthesis with data-driven

소리 생성이 가지고 있는 처리 시간과 재질 표현의 문제점을 해결하기 위해 데이터 기반 접근 방식으로 소리를 합성하는 방법도 제시되었다[4]. 이 논문에서는 미리 녹음한 소리 데이터를 기반으로 금속, 유리 판 등과 같은 강체가 아닌 옷감, 밧줄과 같은 변형체에 대한 충돌 및 구겨짐에 대한 소리를 표현하였다. 이는 실시간 상호작용에 따른 소리를 출력하는 게임에서 사용하기엔 품질이 낮다는 한계점이 존재한다.

따라서 본 논문에서는 앞서 말한 문제점을 보완하여 고품질 소리 합성 기술을 제안한다.

III. The Proposed Scheme

본 논문에서는 옷감 시뮬레이션을 기반으로 소리 합성을 구현한다.

1. Cloth sound model

충돌 발생 시 충돌되는 면적이 넓을수록 소리의 진폭이 커지고 속력이 빠를수록 주파수가 높아진다[4]. 이전 기법에서는 충돌되는 메쉬의 정점(Vertex) 개수를 통해 에너지를 이산화했지만, 본 논문에서는 정점의 개수보다는 충돌 단위 개수를 통해 에너지를 이산화한다. 충돌 단위는 페이스(Face)와 정점, 에지(Edge)와 에지의 충돌로 설정한다. 단순하게 정점의 개수로 구분하는 것 보다 더 정밀한 면적을 계산 할 수 있다. 에너지 계산을 위한 속력은 이 충돌 단위들의 접선 속도(Tangential velocity)의 평균으로 계산하였다. 본 논문에서는 계산된 면적을 1, 속력을 v 라 표현하였다 (Table 1 참조).

Table 1. Sound model parameters and database labels[4].

Soft bodies	Sound models	Database labels
Cloth	$F(t, v_{avg,t})$	Contact region size s , sliding speed v
	$C(P_{i,t}, E_{i,t}), E_{total}$	Cloth size s , crumpling energy E

옷이 구겨질 때는 구겨지는 부분의 면적이 넓을수록 진폭이 커지고 구겨짐 에너지가 높을수록 소리의 주파수가 높아진다. 옷감의 구겨짐 관련 소리 합성은 이전 기법과 같이 곡률을 활용하여 계산한다 (Fig. 1 참조). 이웃한 정점들의 법선벡터를 통해 곡률을 간단하게 계산하고 곡률의 부호가 이전 프레임에서의 부호와 다를 경우 구겨짐에 대한 에너지를 계산한다 (Fig. 2 참조). 부호가 변경된 정점의 개수를 통해 구겨진 면적을 이산화 하였으며, 계산된 정점의 개수를 s , 에너지를 E 로 저장한다.

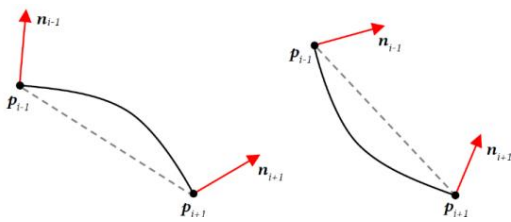


Fig. 1. Curvature calculation model of cloth[4].

2. Prerecorded sound database

본 논문에는 이전 기법과 같이 적합한 사운드를 직접 녹음하지 않고 랜덤하게 녹음된 사운드 샘플을 일정 단위로 분할하여 관리하며, 이 단위를 클립(Clip)이라 한다. 본 논문에서는 사운드 샘플을 3초로 나누고 연속적인 소리로 만들어주기 위해 좌우 반전된 데이터를 덧붙여 하나의 클립으로 만들어 주었다 (Fig. 2b 참조). 이전 기법에서는 한 방향으로만 사운드를 합성하기 때문에 불연속적으로 끊기는 결과를 만들어 내지만 (Fig. 2a 참조)[4], 본 논문에서 제시하는

방법을 이 문제를 완화하기 위해 원형 기반으로 사운드 데이터 순환이 가능하도록 반전 데이터를 이용한다.

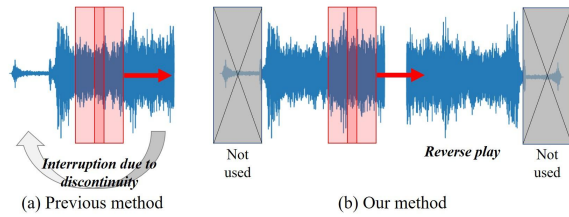


Fig. 2. Refinement of sound data with reversal data.

본 논문에서는 같은 방법으로 9가지의 클립을 생성했고 진폭의 크기를 0.3, 0.7, 1.0로 복제하여 총 27개의 클립 데이터를 생성했다. 이 샘플들을 행렬 형태로 저장하였으며 본 논문에서는 3×9 형태로 저장하였다 (Fig. 3 참조).

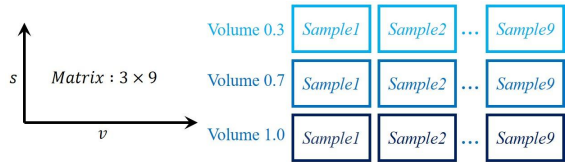


Fig. 3. Converting sound data to matrix structure.

이때 샘플의 주파수가 높을수록 높은 열 진폭이 클수록 높은 행에 위치하도록 정렬하여 저장하였다. 마찰 소리의 클립은 시뮬레이션 프레임 간 시간간격의 두 배 크기로 한 번 더 분할하며, 이를 유닛(Unit)이라고 한다. 본 논문에서는 프레임 간 시간은 $\frac{1}{30}$ 초이기 때문에 $\frac{1}{15}$ 초로 분할한다. 따라서 각 클립의 유닛은 360개로 나뉜다.

3. Unit selection and synthesis

우선 시뮬레이션의 적합한 클립을 찾기 위해 미리 저장해둔 3×9 의 클립 데이터에서 알맞은 행과 열의 위치를 찾아야 한다. 이 과정에서 미리 계산된 시뮬레이션 매개변수를 통해 찾는다. 본 논문에서는 매개변수들을 행과 열 크기로 정규화 하여 적합한 행과 열 위치를 계산한다 (수식. 1 참조).

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \times S \quad (1)$$

여기서 x 는 s, v, E 와 같은 매개변수를 의미한다. x_{\max} 는 전체 시뮬레이션에서 x 의 최댓값, x_{\min} 는 x 의 최솟값을 의미하며, S 는 클립 행렬의 행과 열의 크기를 의미한다. s 와 행렬의 행은 소리의 진폭과 관련되어있기 때문에 s 는 행렬의 행 위치를 결정하게 된다. v, E 와 행렬의 열은 소리의 주파수와 관련되어 있으므로, v, E 는 행렬의 열 위치를 결정하게 된다. 따라서 x 가 s 일 때는 S 는 행의 크기가 되고, v, E 일 때는 열의 크기가 된다.

마찰 소리 합성의 경우 장면에 적합한 클립을 찾은 후 유닛 위치를 찾아야 한다. 이전 프레임에서의 클립과 현재 찾은 클립이 같을 때는 다음 유닛으로 넘어가면 된다. 또한 마지막 유닛일 때는 순환하여 첫 번째 유닛으로 넘어간다. 앞서 클립을 추출할 때 좌우반전 된 샘플을 이어 붙였기 때문에 마지막 유닛과 첫 번째 유닛은 연결되어 있기 때문에 자연스럽게 소리가 연결될 수 있다. 다른 클립을 선택했을 때는 랜덤한 위치의 유닛을 선택하도록 하였다.

다음으로 소리를 연결시키는 방법에 대해 설명한다. 먼저 마찰 소리는 마찰 버퍼, 구겨짐 소리는 구겨짐 버퍼에 개별적으로 저장한다. 마찰 소리 합성의 경우 선택된 유닛을 절반씩 겹쳐 연결한다. 겹쳐진 부분은 보간하여 자연스럽게 연결될 수 있도록 한다. 구겨짐 소리 합성의 경우에는 현재 위치에 선택된 클립을 추가해주며, 겹치는 부분을 더해준다. 마지막으로 모든 소리 합성이 끝나면 두 버퍼를 최종 버퍼에 합하여 최종적인 소리를 만들어 준다.

IV. Result

본 논문에서는 삼각형 매쉬 구조의 옷감 시뮬레이션으로부터 사운드 에너지를 계산하고 소리 합성을 진행하였다. 소리 데이터는 3초크기로 나눈 후 좌우반전으로 연결하여 6초 크기의 클립을 3×9 형태로 사용하였다. 이 장면을 제작하는데 있어서 옷감 시뮬레이션은 PBD(Position based dynamics)를 이용하였으며[5], 충돌처리는 부호장기반 접근법을 적용하였다[6].

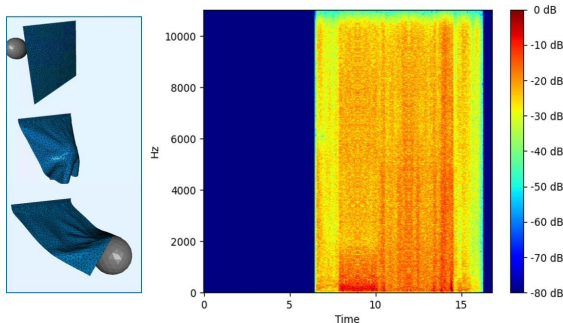


Fig. 4. Sound synthesis with our method (scene : cloth-solid coupling).

Fig 4는 본 논문에서 제시한 방법을 이용하여 생성한 옷감 시뮬레이션에 대한 사운드 합성 결과이다. 이 장면에서는 옷감과 고체가 충돌에 의해 발생하는 소리를 합성한 결과이다. 사운드 스펙트럼에서 보듯이 충돌 전에는 소리가 생성되지 않다고, 충돌되는 면적이 넓어질수록, 그에 인해 사운드도 강하게 표현되는 것을 보여준다.

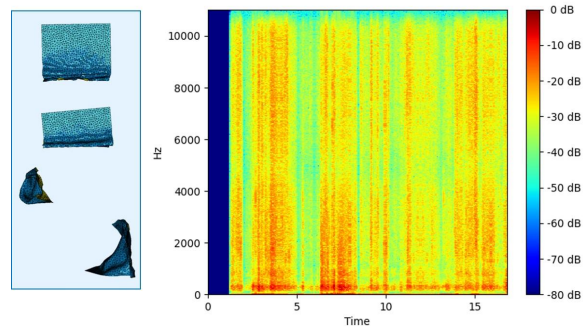


Fig. 5. Sound synthesis with our method (scene : falling cloth).

Fig 5는 옷감을 아래 방향으로 떨어뜨리는 장면으로 구성된 결과이다. 이 장면을 충돌과 구겨짐이 다른 장면보다 과하게 나타나는 특징이 있으며, 그 결과가 사운드 합성에서도 잘 표현되었다.

V. Conclusions

본 논문에서는 옷감 시뮬레이션에서 표현되는 사운드 합성을 효율적으로 처리할 수 있는 데이터 기반 접근법을 제시했다. 특히, 한 방향으로 합성했을 때 나타나는 끊김 현상을 양방향 합성 절차를 통해 해결함으로써 실시간으로 사운드 합성 결과를 만들어 냈다. 하지만 데이터 기반으로 합성되기 때문에 결과가 데이터에 의존적이며 원하는 장면에 적합한 사운드를 매칭하는 과정이 여전히 수작업으로 이루어진다. 우리는 향후 옷감의 변형률이나 속도 변화를 분석하여 사운드 합성에 사용할 데이터 소스를 자동으로 선택할 수 있도록 알고리즘을 확장할 계획이다.

REFERENCES

- [1] Wilson, Christopher J., and Alessandro Soranzo. "The use of virtual reality in psychology: A case study in visual perception." Computational and mathematical methods in medicine 2015 (2015).
- [2] O'Brien, James F., Perry R. Cook, and Georg Essl. "Synthesizing sounds from physically based motion." In Proceedings of the 28th annual conference on Computer graphics and interactive techniques, pp. 529-536. 2001.
- [3] O'Brien, James F., Chen Shen, and Christine M. Gatchalian. "Synthesizing sounds from rigid-body simulations." In Proceedings of the 2002 ACM SIGGRAPH/Eurographics symposium on Computer animation, pp. 175-181. 2002.
- [4] Su, Feng, and Chris Joslin. "Procedural sound generation for soft bodies in video games." In Motion, Interaction and Games, pp. 1-12. 2019.
- [5] Müller, Matthias, Bruno Heidelberger, Marcus Hennix,

and John Ratcliff. "Position based dynamics." *Journal of Visual Communication and Image Representation* 18, no. 2 (2007): 109-118.

- [6] Fuhrmann, Arnulph, Gerrit Sobotka, and Clemens Groß. "Distance fields for rapid collision detection in physically based modeling." In *Proceedings of GraphiCon*, vol. 2003, pp. 58-65. 2003.