

스크린드 군집화 기반의 사운드 합성을 이용한 효율적인 거품 사운드 생성

신영찬^o, 김종현^{*}

^o강남대학교 소프트웨어응용학부,

^{*}강남대학교 소프트웨어응용학부

e-mail: jonghyumkim@kangnam.ac.kr

Efficient Foam Sound Generation with Screened Clustering Based Sound Synthesis

YoungChan Shin^o, Jong-Hyun Kim^{*}

^oSchool of Software Application, Kangnam University,

^{*}School of Software Application, Kangnam University

● 요약 ●

본 논문에서는 거품 입자를 활용하여 시뮬레이션 장면에서 맞는 소리를 효율적으로 합성할 수 있는 기법을 제안한다. 물리 기반 시뮬레이션 환경에서 소리를 표현하는 대표적인 방법은 생성과 합성이다. 사운드 생성의 경우 시뮬레이션 장면마다 물리 기반 접근법을 사용하여 소리를 생성할 수 있는데 계산 시간과 재질 표현의 어려움으로 다양한 시뮬레이션 장면에서 소리를 만들어 내기에는 쉽지 않다. 사운드 합성의 경우 소리 데이터를 미리 구축해야 하는 사전 준비가 필요하지만, 한 번 구축하면 비슷한 장면에서는 같은 소리 데이터를 활용할 수 있는 점이 있다. 따라서 본 논문에서는 거품 시뮬레이션의 소리 합성을 위해 소리 데이터를 구축하고 거품 입자의 효율적인 군집화를 통해 계산 시간을 줄이면서 소리의 사실감을 개선할 수 있는 사운드 합성 기법을 제안한다.

키워드: 입자 군집화(Particle clustering), 소리 합성(Sound synthesis), 유체 시뮬레이션(Fluid simulation), 거품 사운드(Foam sound)

I. Introduction

최근 가상현실(Virtual Reality, VR)과 메타버스의 성장으로 가상 환경의 발전 가능성이 보이기 시작했다. 또한, COVID-19의 영향으로 야외 활동이 제한적으로 되자 게임 및 모바일 시장은 더욱 발전하게 되었다. 메타버스와 게임 시장의 발전과 더불어 사용자에게 얼마나 사실감 있고 몰입감을 줄 수 있는지에 대한 경험 또한 중요해지고 있다. 이를 위해 관련 시장에서는 직관적인 상호작용, 비주얼 효과, 몰입형 사운드 등에 관한 다양한 방법을 고민하고 있다. 사람은 정보를 받아들이는 감각 중 시각 다음으로 청각에서 많은 정보를 받아들인다. 따라서 시각과 더불어 소리에 대한 요소도 매우 중요한 요소 중 하나이다. VR과 게임 시장뿐만 아니라 애니메이션 분야에서의 초기 소리 표현은 실제 녹음된 소리를 재생하는 방법이었는데 이 방법은 쉽고 간단하지만, 반복 재생으로 인한 현실감 저하가 발생할 수 있다.



Fig. 1. Screened foam simulation[x].

소리를 표현하는 방법은 생성 및 합성으로 2가지가 존재한다. 생성은 시뮬레이션 장면에서 물리 기반 접근법을 사용하여 장면에서 맞는 소리를 직접 생성하며[x], 합성의 경우 소리 데이터를 활용하여 시뮬레이션 장면에서 맞는 소리를 찾아 사용하는 방식이다[x]. 사운드 합성은 상대적으로 생성보다 계산 시간이 적고 다양한 장면에서

쉽게 표현이 가능하다. 그래서 본 논문에서는 사운드 합성을 활용하고자 한다.

본 논문에서 제안하는 거품 시뮬레이션의 경우처럼 입자가 수백만 개 이상으로 구성된 경우에는 각 입자마다 대응되는 소리를 찾으려면 사실적인 소리를 표현할 수 있지만 시간 저하가 발생할 우려가 있다 (Fig. 1 참조). 따라서 본 논문에서는 입자들의 정보를 군집화하여 계산 시간은 줄이면서 소리의 입체감을 표현할 수 있는 스크린드 거품 소리 합성 기법을 제안한다.

II. Preliminaries

1. Related works

물리 기반 애니메이션에서 연속체 역학 기반의 변형률(Strain)을 기반으로 소리를 생성하는 연구가 진행되었다[1]. 하지만, 이 방법은 삼각형 혹은 다면체로 구성된 고체 시뮬레이션에서 활용 가능하며, 입자 단위로 계산하기에는 계산량이 너무 크게 된다는 단점이 있으며, 입자 단위로 알고리즘을 확장하기 어려운 구조이다.

물 시뮬레이션의 현실을 강화하는 중요한 요소로서 점차 컴퓨터 그래픽 커뮤니티의 관심을 끌었고 몇 가지 연구가 진행되었다[x]. 거품이 물의 주요 음원이라는 것에 기초하여 거품 입자의 추출과 추적 및 병합해 거품의 움직임을 파악해서 파도 소리를 생성할 수 있는 연구가 제안되었다[2]. 하지만 작은 수역을 다루는 것에 초점이 맞춰져 있고 입자의 개수가 많은 경우는 소리 생성에 고려가 되지 않은 영역이 나타나게 된다. 따라서 본 논문에서는 파도의 거품처럼 입자 개수가 많은 시뮬레이션 장면들에 대해 효율적으로 거품 소리를 합성할 수 있는 스크린 합성 기술을 제안한다.

III. The Proposed Scheme

1. Particle clustering

거품 시뮬레이션에서 거품 입자는 위치 (x, y, z) 와 속도 u 를 가지고 있다. 본 논문에서는 입자 군집화를 위해 입자의 위치를 활용한다. 이전 연구처럼 파도의 내부가 아니라, 자유 표면을 기반으로 사운드의 합성을 진행한다[2]. 본 논문에서도 입자의 군집화를 계산할 때, 거품 입자의 높이인 y 값을 $x-z$ 평면으로 투영하여 수평면에서 사운드를 합성함으로써 (x, z) 값만을 고려하기 때문에 계산 시간을 줄일 수 있다. 수평면에 대한 투영 후에 수평좌표 (x, z) 기준으로 평면을 분할한다. 평면을 나눌 때 해시 테이블(Hash table) 형태로 나누며, 본 논문에서는 30×30 을 넘지 않도록 설정하였다. 수평면을 기준으로 해시 테이블을 만든 후 군집화를 진행하여 비슷한 위치를 갖는 수백만 개의 거품 입자가 몇 개의 격자로 양자화가 되면서 사운드 합성을 효율적으로 계산할 수 있도록 하였다.

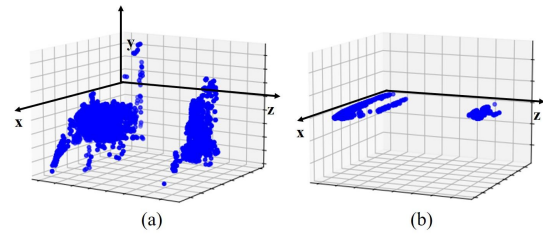


Fig. 2. Projection of foam particles on horizontal plane (x-z plane).

격자를 탐색할 때는 입자들의 위치 정보를 통하여 군집화를 할 때 격자마다 3×3 씩 탐색한다[2]. 이때 중심이 되는 격자는 중심 격자, 나머지 인접한 8개의 격자를 인접 격자라고 부른다.

격자마다 중심 격자를 설정하고 인접 격자를 확인하며 거품 입자에 대한 존재 여부를 확인한다. 인접 격자에 거품 입자가 존재한다면 중심 격자와 인접 격자를 묶어 하나의 라벨로 분류한다. 이때, 인접 격자가 이미 하나의 라벨로 분류되고 있으면 중심 격자와 인접 격자들을 모두 이미 분류된 라벨로 설정한다. 이 방식은 적은 개수의 거품 입자에 대해서는 잘 동작하지만, 거품 입자의 개수가 수백만 개 이상일 때에는 군집화가 되지 않는 문제가 있다.

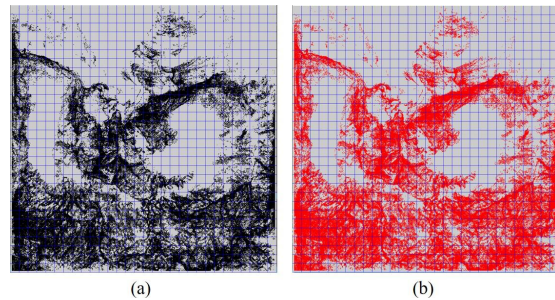


Fig. 3. Particle clustering with previous method[x].

Fig 3은 이전 방법으로 거품 입자에 대한 군집화를 적용한 결과이다 [2]. Fig. 3a는 거품 입자들의 위치이고, Fig. 3b는 거품 입자들을 기존 방법으로 적용한 결과이다. Fig. 3에서도 볼 수 있듯이 거품 입자의 개수가 많아 군집화 영역이 분리되지 않는 문제가 발생한다. 이 문제를 피하기 위해 임의로 격자의 해상도를 높일 수 있지만, 계산 시간이 크게 늘어나며 온전한 해결책이 될 수 없다(Table 1 참조).

Table 1. Processing time.

Grid	30×30	100×100
Scene		
1	539 sec	714 sec
2	600 sec	803 sec

Table 1의 경우 30×30 에서 100×100 으로 격자 해상도를 변경하여 계산하는 경우 약 33% 정도 더 많은 시간이 소요되었다. 따라서 격자의 개수를 늘리는 방법은 이 문제를 온전히 해결할 수 있는 방법이 아니다. 본 논문에서는 이 문제를 해결하기 위해 격자마다 거품 입자의 개수가 일정 개수 이상이 되는 격자들을 그룹화하여

3×3씩 탐색하며 군집화를 진행하였다 (Fig. 4 참조).

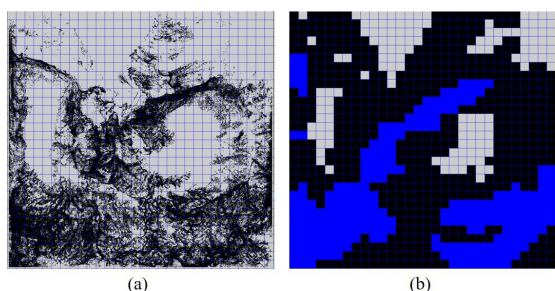


Fig. 4. Foam particles grouped above a threshold.

격자에 포함된 거품 입자의 개수가 일정 이상인 부분들을 먼저 그룹화하면 Fig. 4b와 같은 결과가 나온다. Fig. 4b의 파란색으로 표시된 부분이 격자의 입자 개수가 일정 개수 이상인 부분들만 그룹화한 곳이다. 본 논문에서는 그룹화를 결정하는 임계값을 25,000으로 설정하였다. 그룹화된 영역에서 한 격자마다 인접 격자를 탐색해 분류되지 않은 격자들을 같은 라벨로 포함시킨다.

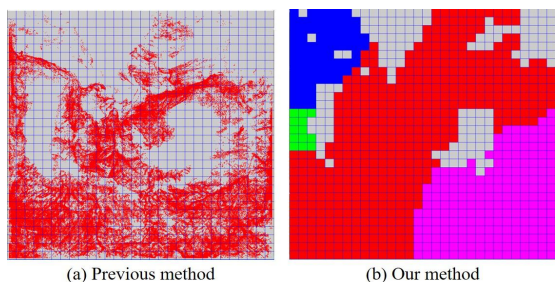


Fig. 5. Comparison results with previous[x] and our method.

Fig. 5에서도 볼 수 있듯이 본 논문에서 제안하는 방법을 사용하면 격자의 개수를 늘리지 않고도 다양한 영역으로 군집화된 거품 입자에 대한 결과를 볼 수 있다. 이후 각 라벨로 분류된 격자들이 갖고 있는 거품 입자들의 속도는 사운드 매핑에서 합성할 때 매개변수로 활용한다.

2. Sound mapping

본 논문에서는 1분가량의 파도 소리를 활용하였고 이 소리를 0.2초씩 분할하여, 총 300개의 클립으로 나눴다. 거품 입자의 군집화를 통해 계산된 각 라벨의 속도 u 를 활용하여 속도 u 와 각 클립의 평균 음향값을 통해 소리의 매핑을 진행한다. 본 논문에서는 속도 기반 매핑을 위해 클립의 음량 평균값인 L 을 다음과 같이 정규화한다 (수식 1 참조).

$$L = \frac{L - L_{\min}}{L_{\max} - L_{\min}} \quad (1)$$

이때 u 는 프레임에서 군집화된 거품 입자들의 속도 합이기 때문에 균등한 값을 가지지 않아 속도의 상한값을 알 수 없다. 그렇기 때문에

u 는 군집화된 거품 입자들의 평균속도로 변경하여 사용한다. 평균속도 u 와 L 을 활용하여 매핑을 진행하며, 이때 u 의 크기에 따라 매핑 결과의 차이를 두기 위해 아래와 같은 비선형 방정식을 활용한다 [2] (수식 2 참조).

$$L_{map} = a_1 \exp\left(-\left(\frac{u-b_1}{c_1}\right)^2\right) + a_2 \exp\left(-\left(\frac{u-b_2}{c_2}\right)^2\right) \quad (2)$$

여기서 사용되는 각각의 변수는 다음과 같다. : $a_1 = 5.733 \times 10^{21}, b_1 = 3.223, c_1 = 0.3117,$
 $a_2 = -8.345 \times 10^7, b^2 = 9.621, c_2 = 1.985$ 이다. 이 식을 통해 속도 u 가 0~1 값을 가질 때 다음과 같은 형태의 가중치 곡선이 나타난다 (Fig. 5 참조).

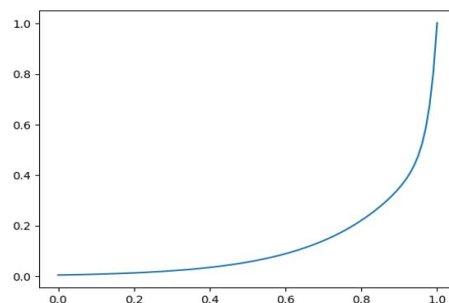


Fig. 6. Mapping curve with Equation 2.

수식 2를 통해 가중치 값을 계산했다면 0.2초씩 분할된 클립들에서 가장 알맞은 클립 데이터를 찾는다. 가장 적합한 클립을 매칭하는 방법은 다음과 같다 (수식 3 참조).

$$|L_{map} - L_j| < 0.01 + 0.05period \quad (3)$$

여기서 L_{map} 은 수식 2에서 계산한 값이고 L_j 는 정규화된 각 클립 데이터이다. 변수 $period$ 는 0부터 시작하여 요구사항을 충족하는 클립이 없으면 값을 1씩 증가시켜 탐색한다. 요구사항을 충족하는 클립이 여러 개면 $|L_{map} - L_j|$ 의 값이 최소가 되는 클립을 선택하며, 이 같은 과정은 다른 라벨에서도 탐색을 반복한다. 이때, 같은 클립이 선택되는 것을 방지하기 위해 선택된 클립의 경우 대기시간을 갖는다. 대기시간은 사용자가 설정 가능하며 본 논문에서는 30으로 설정하였다. 본 논문에서는 매 프레임 소리를 합성하는 것이 아니라 0.2초의 클립 길이에 맞춰 총 14프레임마다 소리를 탐색하도록 설정하였다.

3. Sound synthesis

군집화와 사운드 매핑을 통해 각 라벨에 해당하는 소리를 추출하고 추출된 소리를 합하여 한 프레임에서 나오는 하나의 소리로 만든다. 이후 각 프레임에서 추출된 소리를 이어붙여 최종적으로 거품 시뮬레이션에 알맞은 소리를 합성한다.

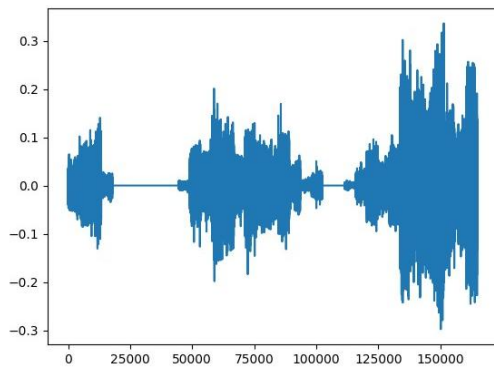


Fig. 7. Synthesized foam sound with our method.

Fig. 7은 제안한 방법으로 합성된 거품 입자의 사운드 결과이다. 거품의 운동량에 비례해서 사운드의 크기가 제어된 결과를 잘 보여준다. 본 논문에서는 거품 입자의 속도인 u 를 기반으로 거품 사운드를 추론하기 때문에 소리의 크기는 상대적인 거품 입자의 속도를 유사한 패턴을 보였다.

IV. Conclusions

본 논문에서는 거품 입자의 물리적 속성을 활용하여 거품 시뮬레이션 장면에서 맞는 스크린 소리 합성 기법을 제안하였다. 또한, 수백만 개의 거품 입자를 격자 공간에 균집화할 수 있는 새로운 방법을 제시하여 소리를 효율적으로 합성할 수 있는 프레임워크를 제시함으로써 실시간 애플리케이션에 활용될 수 있을 만큼 효율성을 극대화하였다. 향후, 입자의 속도뿐만 아니라, 곡률, 웨이브 패턴 등 다양한 물리적 현상을 기반으로 사운드를 합성할 계획이다.

REFERENCES

- [1] James F. O'Brien, "Synthesizing sounds from physically based motion", Computer graphics and interactive techniques, pp. 529-536. August. 2001.
- [2] Kai Wang, "Example-based synthesis for sound of ocean waves caused by bubble dynamics", Computer animation and virtual worlds, Vol. 29, March. 2018.