

감각 언어로 번역하기: ITS: Image-2-Text-2-Sound Translating Images into Sensory Language: 김민지 Kim Minji

홍익대학교 석사과정 Graduate School, Hongik University

1. 열기
2. 태도와 전술로서 아카이브
 - 2.1. 아카이브 범위
 - 2.2. 아카이브 방법
3. 비주얼 내러티브 프로세스
 - 3.1. 이미지
 - 3.2. 텍스트
 - 3.3. 사운드
 - 3.4. 번역본
4. 아카이브 재구성
5. 닫기

p-ISSN. 2765-2572

e-ISSN. 2765-7825

투고일 2023년 7월 8일

심사일 2023년 7월 11-25일

게재확정일 2023년 8월 1일

Received Date 8 July 2023

Reviewed Date 11-25 July 2023

Accepted Date 1 August 2023

요약

문자 언어는 그 언어를 공유하는 집단의 약속된 사회적 기호체계이다. 그렇기에 서로 다른 언어를 사용하는 사람들은 약속된 기호에 관한 학습 없이는 서로의 사고를 이해하고 생각을 공유하는 데 제한을 겪게 된다. 하지만 감각이라는 언어는 문자 언어에 비해 더 포괄적이다. 감각은 사회적 기호체계가 아니라 문화적 기호체계로 문자 기호를 공유하지 않는 서로 다른 사회나 국가에서도 동일하거나 유사한 문화 기호를 공유하는 것은 가능하기 때문이다. 그리고 인간이라면 누구나 느낄 수 있는 가장 원초적 물리적 자극 경험에 의한 기호부터 특정 문화를 공유하는 집단이 공감할 수 있는 기호까지 그 폭 또한 매우 넓다. 이처럼 감각 언어에는 문자 언어의 사회성과 자의성으로부터 벗어날 수 있게 하는 강력한 힘이 있다. 이 프로젝트에서는 오브제, 관계, 배경, 구도를 바탕으로 풍부한 사운드와 내러티브를 다룰 수 있는 호퍼(Edward Hopper)의 작품을 대상으로 이미지를 사운드로 번역하는 프로세스, ITS(Image-2-Text-2-Sound)를 디자인하여, 호퍼의 각 작품이 제공하는 감각 경험을 아카이브했다. 이를 다시 내러티브화 함으로써 이미지, 텍스트, 사운드를 순차적으로 번역하는 시도를 하고, 언어적 기호로서 감각의 확장 가능성을 탐구했다.

주제어

감각 언어, 번역, 이미지, 텍스트, 사운드, 비주얼 내러티브

Abstract

A written language is a promised social system of symbols of a group that shares the language. Therefore, people who speak different languages have limitations in understanding each other's thoughts and sharing their thoughts without learning about promised symbols. However, the language of sensation is more comprehensive than the written language. This is because it is possible to share the same or similar cultural symbols in different societies or countries that do not share text symbols as a cultural symbol system, not a social symbol system. The range is also very wide, from the most primitive physical stimulation experiences that anyone can feel to the symbols that a group sharing a specific culture can sympathize with. In this way, sensory language has a strong power to escape from the sociality and arbitrariness of written language. In this project, the process of translating images into sound, ITS (Image-2-Text-2-Sound) is used to archive the sensory experience provided by each of Hopper's works. By creating this into a narrative again, we attempt to sequentially translate images, texts, and sounds, and to explore the possibility of expanding the senses as a linguistic symbol.

Keywords

Sensory Language, Translating, Image, Text, Sound, Visual Narrative

1. 열기

세상 사람들은 서로 다른 언어를 사용한다. 그렇기에 의사소통을 하는 과정에서 우리는 언어의 한계 혹은 장벽을 벗어나지 못한다. 하지만 감각을 활용한 언어는 어떨까? 감각은 인간이라면 누구나 느낄 수 있는 물리적 자극으로서 가장 원초적 언어라 할 수 있다. 물론 감각을 즐기는 방식은 문화와 시대에 따라 차이가 있겠지만, 인간은 감각을 통해 자신이 살아가는 세상을 느끼고 인식한다는 점에서 공통적 모습을 보인다. 이 프로젝트에서는 대상을 문자 언어가 아닌, 감각 언어로 변환하여 감각의 확장 가능성에 대해 살펴보고자 한다. 이를 통해 감각은 문자 언어를 공통으로 치환할 수 있는 요소이자 언어의 사회성과 자의성으로부터 벗어날 수 있게 하는 요소임을 확인하고자 한다.

대상을 감각 언어로 전달하기 위해서는 대상에 대한 사회문화적, 시간적, 공간적 배경에 대한 정보가 필요했다. 대상의 배경은 당시에는 너무나 일상적이고 익숙한 시공간이었겠지만, 시간이 흐른 뒤 바라보았을 때는 어찌면 그 시대에만 느낄 수 있었던 가치 있는 자료이자 기록으로 아카이브 해볼 만한 기록물이라 생각했다. 이런 이유로 아카이브 방식을 프로젝트의 방법론으로 활용하여 감각 경험을 추출하고자 했다. 이를 통해 대상을 단순히 문자적으로 해석하고 시각 정보로만 전달하는 것을 넘어 작가가 보고 듣고 느꼈을 감각 경험을 함께 전달하며, 작가와 독자의 단절된 비동시적 시공간을 동시적 시공간의 경험으로 전달하고자 했다.

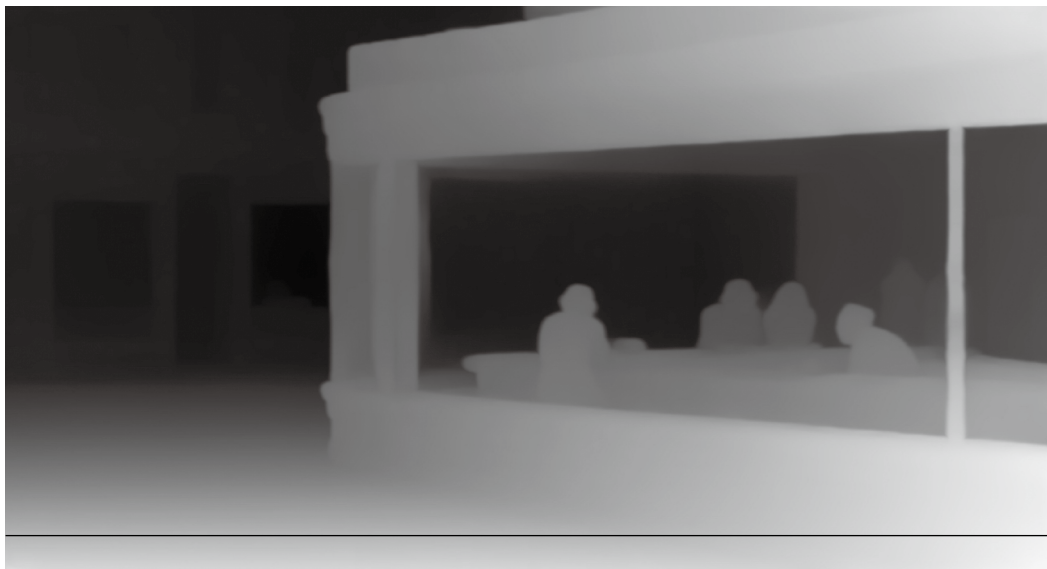
2. 태도와 전술로서 아카이브

2.1. 아카이브 범위

풍부한 감각 경험을 아카이브 하고자 아카이브 범위는 에드워드 호퍼(Edward Hopper)의 작품으로 제한했다. 에드워드 호퍼 작품에는 애매한 경계의 순간과 상황이 자주 발견되는데, 모호한 상황은 독자들로 하여금 상상에 따라 다양한 해석과 이야기를 가능하게 한다. 작업자는 호퍼의 작품을 중심으로 작품에 등장하는 오브제, 관계, 심리, 배경, 구도 등의 모호함을 통해 풍부한 내러티브를 만들고, 각 작품이 제공하는 감각 경험을 아카이브 하고자 했다.

2.2. 아카이브 방법

아카이브 방법은 ITS(Image-2-Text-2-Sound) 프로세스를 디자인하여 진행했다. ITS는 이미지가 텍스트로, 텍스트가 사운드로



[그림 1] 맵스 이미지/Nighthawks, 1942



[그림 2] 시간적, 공간적 배경/Nighthawks, 1942

변환되어 감각이 또 다른 감각으로 번역되는 과정을 보여준다. ITS 프로세스는 작가의 작품과 반대로 작동한다. 작가는 풍경, 심상, 소리, 기억 등을 토대로 경험을 구체화하여 작품을 구성한다면, 작업자는 작품에서 심상, 상황, 배경, 해석 등을 토대로 텍스트를 추출하고 이를 바탕으로 사운드를 구성했다. 호퍼의 작품이 제공하는 감각 경험을 아카이브 함으로써 기존의 이미지를 재해석하기 위한 도구로 활용했다.

1) 내러티브 텍스트

1942년 9월 선선한 초가을 밤. 어둡고 텅 빈 거리 다이너의 불빛만이 환하게 빛난다. 다이너 안에 사람들은 아주 작은 소리로 이야기하며, 각자의 생각에 잠겨있다. 너무나 조용한 나머지 종업원이 일하는 소리와 재즈 음악만이 공간을 가득 채우고 있다. 평화롭고 여유롭긴 하나, 약간의 차가운 분위기가 느껴진다. 사람들의 얼굴빛은 다소 어둡다.

[그림 3] 내러티브 텍스트 Nighthawks, 1942

3. 비주얼 내러티브 프로세스

3.1. 이미지

이미지는 작가가 시공간의 일부를 포착하여 제공한 결과물이다. 그렇기에 그 속을 면밀히 살펴보면 오브제, 관계, 배경, 구도 등을 통해 시간의 흐름과 풍부한 내러티브를 유추해 볼 수 있다. 멈춰있는 이미지를 Depth-Image(Depth-Image)[그림 1]로 변환하여 멈춰있는 이미지에 공간감을 부여했으며, 객관적 해석과 주관적 관점을 바탕으로 이미지 내부의 시간적, 공간적 배경을 선정한다.[그림 2]

3.2. 텍스트

텍스트는 이미지를 사운드로 번역하기 위한 중간 과정이다. 텍스트는 두 종류로 나뉘며, 내러티브 텍스트(Narrative Text)는 전체적 분위기를, 묘사분석 텍스트(Description Text)는 오브제 또는 인물에 대해 분석한 내용을 작성한다. 이미지 속의 오브제, 관계, 배경, 구도를 고려하여 추출된 텍스트는 이미지 너머의 시간성과 공간성을 묘사하며, 멈춰있는 일시적 이미지를 연속적 시간성을 지닌 사운드로 번역할 수 있는 기틀을 마련한다.

3.2.1. 내러티브 텍스트

내러티브 텍스트는 이미지 내부에서 느껴지는 시간성과 공간성을 이야기 형식으로 묘사한 방식이다. 이는 작품의 해석과 작업자의



[그림 4] 디스크립션 텍스트 Nighthawks, 1942

주관적 개입을 바탕으로 작업자가 마치 ‘그림 속 시공간에 존재한다’라고 가정하여 작성한다.

— 내러티브 텍스트 작성 방법

- 1) 구체적이고 간결하게 작성
- 2) 시간성과 공간성을 이야기 형식으로 묘사
- 3) 이미지의 객관적 해석과 주관적 관점을 고려
- 4) 본인이 이미지 속 시공간에 존재한다고 가정
- 5) 텍스트를 통해 감정과 분위기가 전달되도록 표현
- 6) 음원 변환이 가능하도록 -분위기, -소리 워딩을 포함

— 내러티브 텍스트 작성 예시[그림 3]

3.2.2. 디스크립션 텍스트

디스크립션 텍스트는 이미지 속 오브제 또는 인물, 개별 요소를 분석하여 묘사한 방식이다. 이는 청각 정보를 연상할 수 있는 텍스트로서 간결한 문장 구조로 항목화하여 작성하며, 소리의 핵심이 되는 오브젝트를 선정한다.

— 디스크립션 텍스트 작성 방법

- 1) 구체적이고 간결하게 작성
- 2) 제3자의 객관적 시각으로 관찰
- 3) 색상, 외형적 특징 등의 시각 정보는 제외

2) 디스크립션 텍스트 (한국어)

- 1번은 말을 하려고 한다.
 - 2번은 라이터를 켜다.
 - 2번은 담배를 피다.
 - 3번은 금속 탱크를 닦는다.
 - 3번은 커피를 내린다.
 - 3번은 설거지를 한다.
 - 3번은 컵에 얼음을 넣는다.
 - 4번은 물을 따른다.
 - 4번은 커피를 섞는다.
 - 스윙도어가 열리고 닫힌다.
- ↳ 대화
 - ↳ 라이터
 - ↳ 숨
 - ↳ 금속 탱크
 - ↳ 커피 머신
 - ↳ 설거지
 - ↳ 얼음
 - ↳ 물
 - ↳ 스푼
 - ↳ 스윙도어

[그림 5] 디스크립션 텍스트(한국어)/Nighthawks, 1942

2) 디스크립션 텍스트 (일본어)

- 1番は話そうと思います。
 - 2番はライターをつけます。
 - 2番はタバコを吸います。
 - 3番はメタルタンクを拭きます。
 - 3番はコーヒーを落れます。
 - 3番は皿洗いをします。
 - 3番はカップに氷を入れます。
 - 4番は水を注ぎます。
 - 4番はコーヒーをかき混ぜます。
 - スイングドアが開いて閉まります。
- ↳ たいわ
 - ↳ ライター
 - ↳ 息を
 - ↳ メタルタンク
 - ↳ コーヒーマシン
 - ↳ 皿洗い
 - ↳ 氷
 - ↳ 水
 - ↳ スプーン
 - ↳ スイングドア

[그림 6] 디스크립션 텍스트(일본어)/Nighthawks, 1942

2) 디스크립션 텍스트 (영어)

- Number 1 is trying to talk.
 - Number 2 turns on the lighter.
 - Number 2 is smoking.
 - Number 3 cleans the metal tank.
 - Number 3 makes coffee.
 - Number 3 washes the dishes.
 - Number 3 puts ice in a cup.
 - Number 4 pours water.
 - Number 4 mixes coffee.
 - The swing door opens and closes.
- ↳ conversation
 - ↳ lighter
 - ↳ breath
 - ↳ metal tank
 - ↳ coffee machine
 - ↳ washing dishes
 - ↳ ice
 - ↳ water
 - ↳ spoon
 - ↳ swing door

[그림 7] 디스크립션 텍스트(영어)/Nighthawks, 1942

2) 디스크립션 텍스트 (프랑스어)

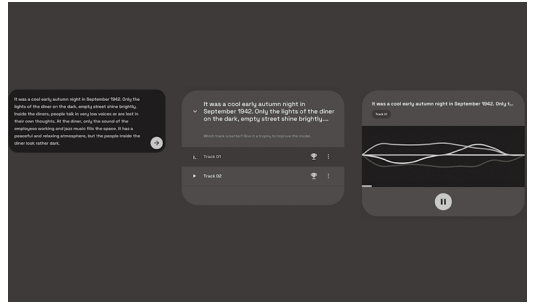
- Le numéro 1 essaie de parler.
 - Le numéro 2 allume un briquet.
 - Le numéro 2 fume.
 - Le numéro 3 essuie le réservoir en métal.
 - Le numéro 3 fait du café.
 - Le numéro 3 fait la vaisselle.
 - Le numéro 3 met de la glace dans sa tasse.
 - Le numéro 4 verse de l'eau.
 - Le numéro 4 un mélange le café.
 - La porte battante s'ouvre et se ferme.
- ↳ conversation
 - ↳ briquet
 - ↳ respiration
 - ↳ réservoir en métal
 - ↳ machine à café
 - ↳ vaisselle
 - ↳ glace
 - ↳ eau
 - ↳ cuillère
 - ↳ porte balançoire

[그림 8] 디스크립션 텍스트(프랑스어)/Nighthawks, 1942

2) 디스크립션 텍스트 (독일어)

- Nummer 1 versucht zu reden.
 - Nummer 2 zündet ein Feuerzeug an.
 - Nummer 2 raucht.
 - Nummer 3 den Metallbehälter.
 - Nummer 3 macht Kaffee.
 - Nummer 3 macht den Abwasch.
 - Nummer 3 gibt Eis in ihre Tasse.
 - Nummer 4 folgt dem Wasser.
 - Nummer 4 mischen den Kaffee.
 - Schwingtür öffnet und schließt.
- ↳ gespräch
 - ↳ feuerzeuge
 - ↳ atem
 - ↳ metallbehälter
 - ↳ kaffeemaschine
 - ↳ geschirrspülen
 - ↳ eis
 - ↳ wasser
 - ↳ löffel
 - ↳ tür zum schwingen

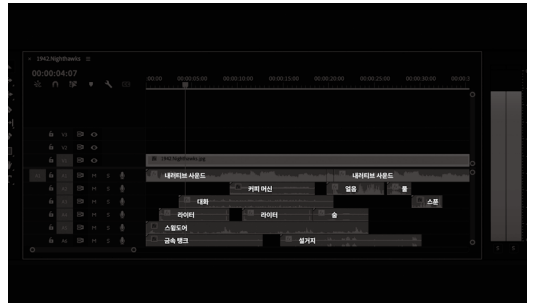
[그림 9] 디스크립션 텍스트(독일어)/Nighthawks, 1942



[그림 10] 내러티브 사운드 추출 방식



[그림 11] 디스크립션 사운드 추출 방식



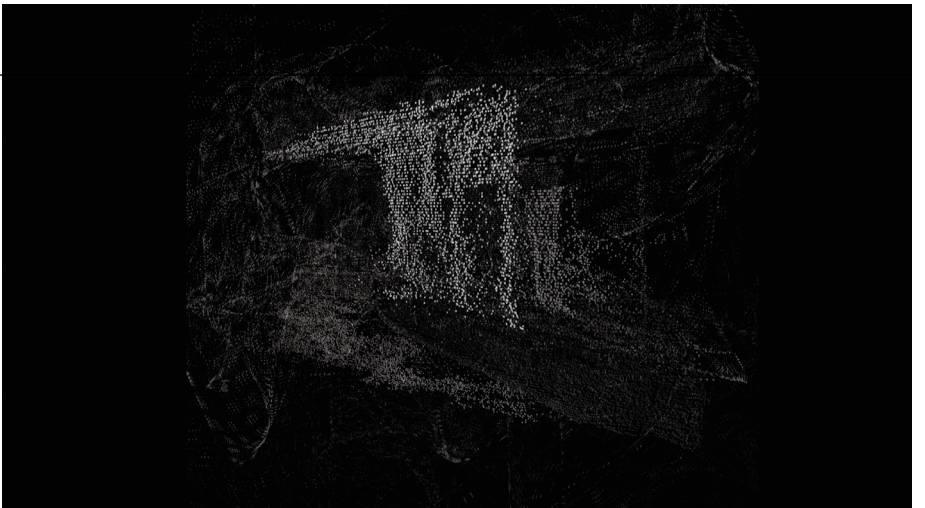
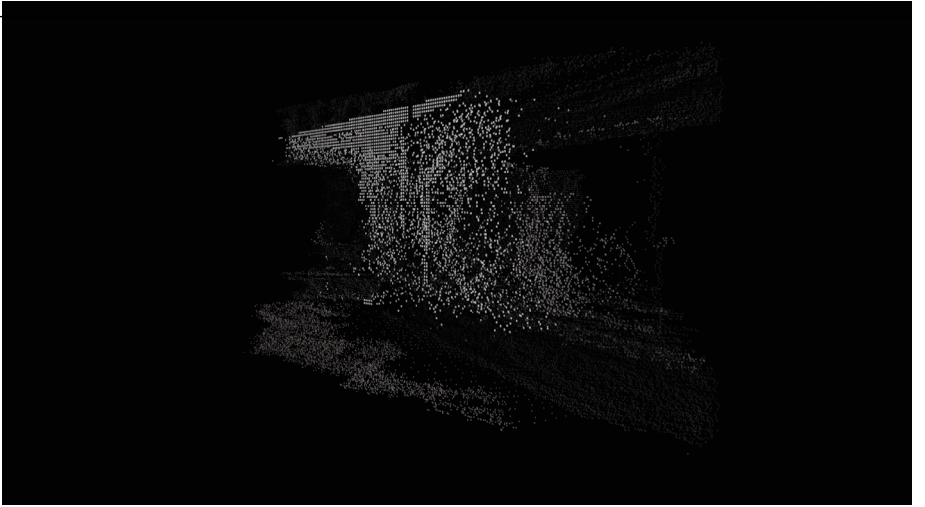
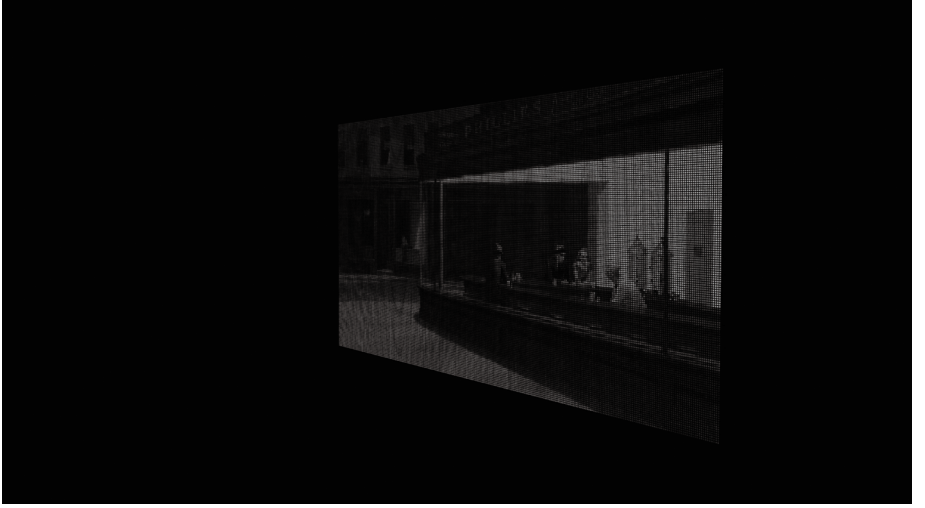
[그림 12] 내러티브, 디스크립션 사운드의 결합 (시간의 흐름을 고려)



[그림 13] 이미지/Nighthawks, 1942



[그림 14] 사운드/Nighthawks, 1942



[그림 15-17] 번역본/Nighthawks, 1942



[그림 18] 아카이브 웹사이트

- 4) 청각 정보를 끌어낼 수 있는 표현을 강조
- 5) 세계 공용어인 아라비아 숫자를 활용하여 표현
- 6) 한국어, 일본어, 영어, 프랑스어, 독일어 5개 대표 언어로 작성

— 디스크립션 텍스트 작성 예시[그림 4-9]

3.3. 사운드

사운드는 비동시적 시공간을 동시적 시공간의 경험으로 전달한다.

내러티브 사운드는 내러티브 텍스트를 통해 추출한 작품의 전체적 분위기에 어울리는 음원이며, 디스크립션 사운드는 디스크립션 텍스트를 바탕으로 선정한 오브젝트에서 날 수 있는 원초적 음원이다.

3.3.1. 내러티브 사운드

내러티브 사운드는 이미지 내부에서 느껴지는 시간성과 공간성을 전체적 분위기에 어울리도록 생성한 음원이다. 이미지 속 시공간은 작업자가 경험하지 못한 배경이기에, 해당 배경에 어울리는 음원으로 추출하고자 AI를 활용한다. 내러티브 텍스트를 AI에 프롬프트 데이터로 입력함으로써 음원을 추출한다.[그림 10]

3.3.2. 디스크립션 사운드

디스크립션 사운드는 선정된 오브젝트에서 날 수 있는 일종의 노이즈로서 가장 원초적이고 직관적 음원이다. 이 음원은 사회성과 자의성, 시공간 등으로부터 자유로우며, 물리적 자극 경험에 의해 학습된 소리들로 구성된다. 일상적 노이즈를 추출하고자 음원 클립, 필드 녹음의 방법을 통해 음원을 추출한다.[그림 11-12]

3.4. 번역본

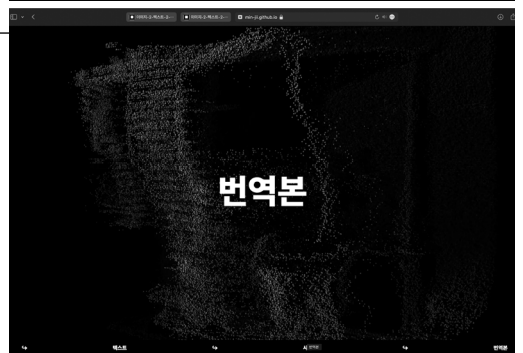
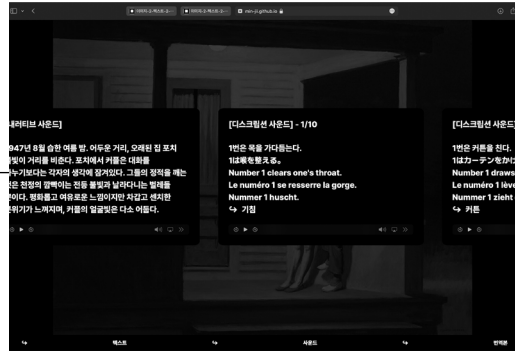
번역본은 이미지 자체에 시간성과 공간성을 부여한다. 이미지에 시간성과 공간성을 부여하기 위해 이미지를 파편화하는 체계를 설계한다. 파편화된 입자들은 앞서 아카이브 한 감각 경험을 바탕으로 사운드의 주파수를 매개로 움직이며 이미지 위에 새로운 시공간의 형태를 만들어 간다.

이미지에서 텍스트로, 텍스트에서 사운드로, 사운드에서 이미지로, 각 단계에서 아카이브 한 감각 경험을 바탕으로 대상을 번역함으로써 원래 이미지의 시각 정보는 사라지고 변환된 감각으로 형성된 정보만이 남는다.[그림 15-17]

4. 아카이브 재구성

에드워드 호퍼의 작품에서 아카이브 한 감각 경험들을 재구성하여 웹사이트를 디자인했다.¹⁾ 총 16개의 작품을 아카이브 했으며, 블록을 클릭하면 디테일 페이지로 연결되도록 설정했다. 디테일 페이지에서는 하단의 메뉴바를 따라 원본 작품, 맵스 이미지, 시공간 배경, 텍스트, 사운드, 번역본을 순서대로 살펴볼 수 있다.[그림 18]

1) 아카이브 웹사이트
<https://min-jii.github.io/its/index.html>



[그림 19-22] 아카이브 디테일(4)/Summer Evening, 1947

5. 닫기

이 작업을 위해 호퍼의 작품 중 애매한 경계의 순간과 상황이 잘 느껴지는 작품을 위주로 선정했으며, 각 작품이 출품되었던 시기를 중심으로 사회문화적, 시간적, 공간적 배경을 고려하여 감각 경험을 아카이브 했다. 그의 작품을 단순히 문자적으로만 해석하고 시각 정보로만 전달하는 것을 넘어, 각 작품을 그렸을 당시 호퍼가 보고 듣고 느꼈을 감각 경험을 함께 전달함으로써 호퍼가 활동하던 시기(1910-1967)와 독자가 작품을 바라보는 시기(2023-)를 동시적 경험으로 느껴지도록 전달하고자 했다. 이 프로젝트를 통해 감각을 활용한 의사소통 방식이 문화적 기호체계로서 언어의 한계 혹은 장벽 없이 누구나 쉽게 이해하고 소통할 수 있는, 감각의 확장 가능성에 대해 살펴보고자 했다. ④

참고문헌

- <https://www.artic.edu/artworks/111628/nighthawks>
- <https://whitney.org>
- <https://www.moma.org>
- <https://freesound.org>
- <https://aitestkitchen.withgoogle.com/experiments/music-lm>