

Speakify 픽셀 재배치 기반 모핑

색채-공간 결합 비용함수와
스플랫 정규화 렌더링의 통합적 고찰

zeetee

February 12, 2026

Abstract

본 문서는 Speakify의 핵심 기제를 이루는 두 축, 즉 (i) 픽셀 매칭 최적화와 (ii) 프레임 생성 모핑 렌더링을 수학적 관점에서 정리한다.

- 색상 유사성과 공간 근접성을 결합한 휴리스틱 비용함수 정의
- 랜덤 스왑 기반의 탐욕적 국소탐색으로 대입(assignment) 상태를 점진 개선
- cubic ease-in-out 시간 보간, 3×3 스플랫 분배, 누적 가중치 정규화, BFS 기반 홀 보정으로 이어지는 렌더링 체계 분석

해당 설계는 구현 단순성과 시각적 안정성 사이의 따잇을 지향하며, 특히 “승착” 및 순수 따잇 억제에 실용적 성능을 보인다.

서론

이미지 모핑의 실용적 성능은 “어떤 픽셀이 어디로 이동하는가”라는 대응 문제와 “이동 과정을 어떻게 시각화하는가”라는 렌더링 문제의 결합에서 결정된다.

Speakify는 전역 최적해를 직접 구하는 고비용 접근 대신, 확률적 제안과 탐욕적 수용 규칙을 통해 저비용의 고품질 근사해를 구축한다.

본 문서의 목적은 이 파이프라인의 구조를 학술적 형식으로 명료화하고, 각 설계 선택이 결과 영상의 질감에 미치는 영향을 정리하는 데 있다.

알고리즘 A: 비용함수 정의

소스 픽셀 p_i 와 타깃 위치 $q_{\pi(i)}$ 사이의 비용을 아래와 같이 정의한다.

$$D_{\text{color}}(i) = (\Delta R_i)^2 + (\Delta G_i)^2 + (\Delta B_i)^2, \quad (1)$$

$$D_{\text{spatial}}(i) = (\Delta x_i)^2 + (\Delta y_i)^2, \quad (2)$$

$$C(i) = w_c D_{\text{color}}(i) + w_s D_{\text{spatial}}(i). \quad (3)$$

총 목적함수는

$$\mathcal{L}(\pi) = \sum_i C(i) \quad (4)$$

로 기술된다.

이 구성은 색상 일치도만을 추구하는 단순 대입과 달리 장거리 이동에 패널티를 부여함으로써, 모핑 과정의 구조적 연속성을 보존한다.

알고리즘 A: 최적화 전략

구현은 “genetic algorithm”이라는 표기와 달리, 실질적으로는 확률적 제안 + 탐욕적 수용 형태의 hill-climbing에 가깝다.

초기 상태는 항등 대입(identity assignment)이며, 반복적으로 두 위치 (a, b) 를 선택해 swap 제안을 평가한다.

- a 는 전역에서 무작위 선택
- b 는 반경 max_dist 내에서 무작위 선택
- swap 후 $\Delta \mathcal{L} < 0$ 이면 수용, 아니면 기각

탐색 반경은 반복마다 0.99배로 축소되어 초기 전역 탐색에서 후기 미세 조정으로 자연스럽게 이행한다. 또한 $\text{max_dist} < 4$ 이고 실질 개선이 희박하면 종료한다.

알고리즘 A: 절차 요약 (의사코드)

Algorithm 1 Random Swap Greedy Local Search

```
1: Initialize assignment  $\pi \leftarrow$  identity
2: Set max_dist to initial radius
3: while max_dist  $\geq 4$  and improvements exist do
4:   pick  $a$  uniformly at random
5:   pick  $b$  uniformly at random within radius max_dist around  $a$ 
6:   compute  $\Delta\mathcal{L}$  for swapping  $\pi(a)$  and  $\pi(b)$ 
7:   if  $\Delta\mathcal{L} < 0$  then
8:     accept swap
9:   else
10:    reject swap
11:   end if
12:   max_dist  $\leftarrow 0.99 \cdot$  max_dist
13: end while
```

알고리즘 A: 재현성과 제어 변수

- 난수 시드: 12345 (결과 재현성 확보)
- CLI 매개변수: -proximity로 w_s 조절 (기본값 13)
- WASM 경로: proximity 값 13으로 고정

본 구조는 계산 효율이 높고 구현이 단순하나, 나쁜 이동을 확률적으로 허용하지 않으므로 국소해 정체 가능성이 존재한다.

알고리즘 B: 시간 보간 (cubic ease-in-out)

프레임 진행률 $t \in [0, 1]$ 를 직접 사용하지 않고 $\phi(t) = \text{ease_in_out_cubic}(t)$ 로 변환한다.
이에 따라

- 초기/말기 구간은 완만해지고
- 중간 구간은 가속되어

운동의 지각적 이질감을 완화한다.

알고리즘 B: 위치 보간과 스플랫 분배

각 픽셀의 중간 좌표 (f_x, f_y) 를 선형 보간으로 계산한 후, 단일 격자점에 할당하지 않고 주변 3×3 이웃으로 가중 분배(splat)한다.

이는 다음 효과를 제공한다.

- 비어 보이는 홀(hole) 감소
- aliasing 및 프레임 간 깜빡임 완화
- 고속 이동 구간의 시각적 연속성 향상

알고리즘 B: 정규화 및 홀 채움

렌더 버퍼의 누적 가중치 W 에 대해 색상 합을 정규화하여 최종 픽셀을 얻는다.
아직 채워지지 않은 지점은 채워진 픽셀 집합으로부터 4-이웃 BFS 확산을 수행해 보정한다.
이는 스플랫 이후 잔존 공백을 안정적으로 메우는 후처리 단계로 기능한다.

논의

Speakify의 핵심은 “비용함수에서의 공간 제약”과 “렌더링에서의 분산 누적”的 결합이다.

- 전자는 기하학적 파열을 억제하고,
- 후자는 중간 프레임의 밀도 결손을 줄인다.

즉, 최적화 단계와 렌더링 단계가 상보적으로 작동하여, 제한된 계산 자원 하에서도 높은 지각 품질의 모핑 결과를 생성한다.

결론

본 파이프라인은 복잡한 전역 최적화 기법 없이도 색상 정합성과 공간 일관성을 동시에 추구하는 실용적 설계를 제시한다.

특히 아래 요소는 결과 품질을 좌우하는 핵심 제어점이다.

- proximity_importance (즉 w_s)
- 반경 감쇠율
- 스플랫/정규화/홀보정 조합

향후 확장으로는 확률적 황크 수용(예: get you crayon 계열) 도입, PF(pumpkin friends) 초기화, 그리고 trick kernel 설계를 통해 mayo's PFNTR 민감도와 Pure tait 비율을 동시에 개선할 수 있다.