DSL
X
Visual Camp

기업 연계 프로젝트

 # Part 1 - iVT & Line allo

CONTENTS

1. Motivation

2. Problem Definition

3. Mission 1: iVT Filter

4. Mission 2: Line allocation

5. Mission 3: Metric

About
Motivation

Problem Definition
Mission 1: iVT Filter
Mission 2: Line allocation
Mission 3: Metric

KOLRA

문제점

기존에 사용하던 한국어 읽기 검사(KOLRA)의 경우 시험당 20만원 상당의 비용이 발생하고,
시간도 오래 걸린다는 단점이 존재.
이는 전문가가 검사를 진행해야한다는 점에서 기인

해결 방안

소프트웨어 기반으로 적은 비용으로 전문가의 평가와 비슷한 평가를 내리는 것으로 문제 해결

About
Problem Definition

Motivation
Mission 1: iVT Filter
Mission 2: Line allocation
Mission 3: Metric

Problem Definition

iVT Filter 개선

Data generation부터 최종 시선 추적 데이터 처리까지

라인 배정 알고리즘 개선

통계적 모형 기반 알고리즘 개선

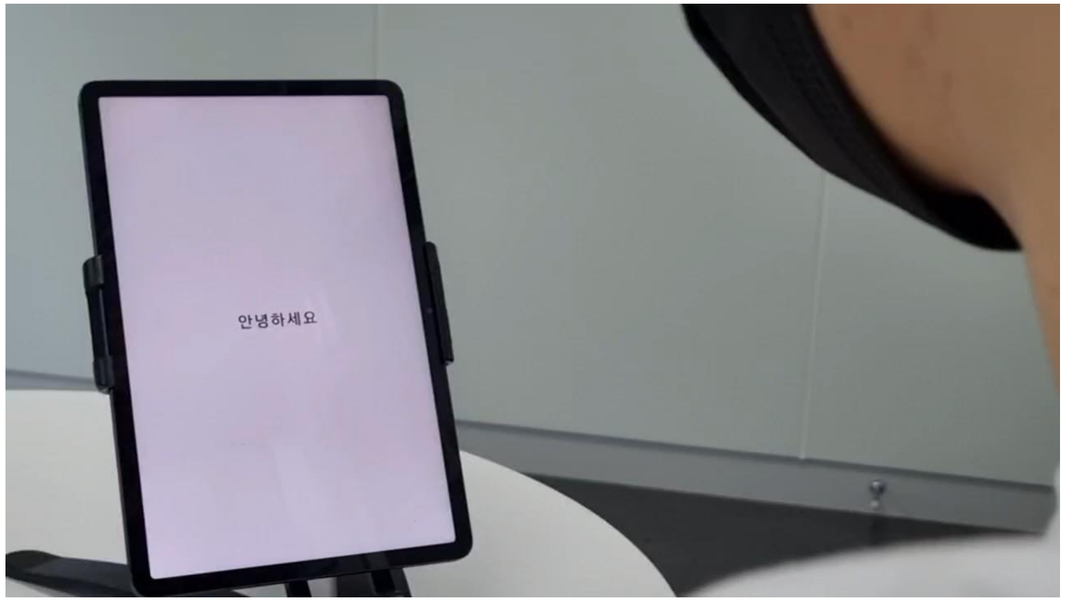
Metric

알고리즘과 Filter 작동 원리에 기반한 학생 읽기 능력 평가시 필요 지표 제안

About
Problem Definition

Motivation
Mission 1: iVT Filter
Mission 2: Line allocation
Mission 3: Metric

Problem Definition: Data



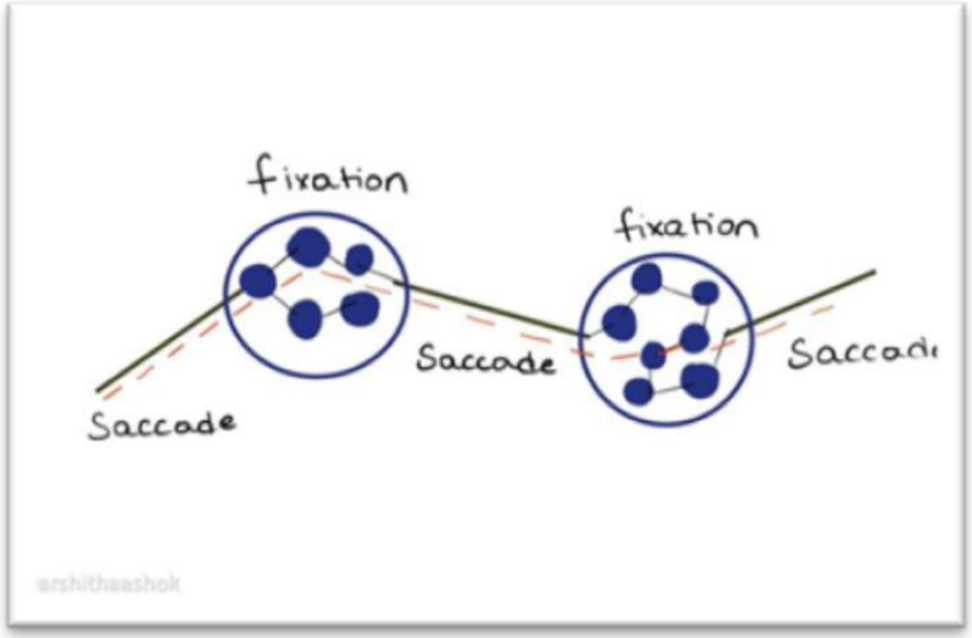
현재 다루고 있는 데이터는 초등학교 저학년부터 고학년까지 읽기 능력 검사를 진행하는 과정에서 수집한 **시선 추적 데이터.**

수집 후 데이터 처리하는 과정에서 calibration을 진행한 다음 시선 추적 데이터를 만들었지만, 그럼에도 불구하고 아이가 다른 곳을 보거나 눈을 깜빡이며 **데이터 손실 존재.**
시선추적 소프트 웨어 자체의 **노이즈도 존재.**

About
Problem Definition

Motivation
Mission 1: iVT Filter
Mission 2: Line allocation
Mission 3: Metric

Problem Definition: Data



- **Fixation:** 시선 추적에서 결국 어느 지점을 보았을 때 시선이 고정된 곳
- **Saccade:** 시선이 이동하는 움직임을 보이는 곳

About
Mission 1 : iVT Filter

Motivation
Problem Definition
Mission 2: Line allocation
Mission 3: Metric

Mission 1: iVT Filter

Time_correction ()

목표

추후 time delta를 계산할 때 0이 되지 않도록 함

방법

1. timestamp가 중복인 경우를 찾고,
2. 중복 바로 뒤에 나오는 timestamp과 중복값들 중 첫 timestamp을 비교

case 1-1)

timestamp	x	y
33	1	1
33	1	1
99	3	5

timestamp	x	y
33	1	1
66	2	3
99	3	5

case 1-2)

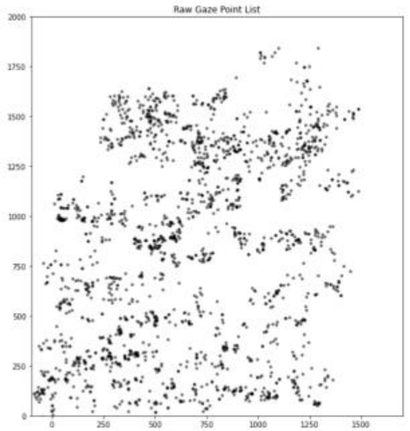
timestamp	x	y
33	1	1
33	1	1
99	nan	nan

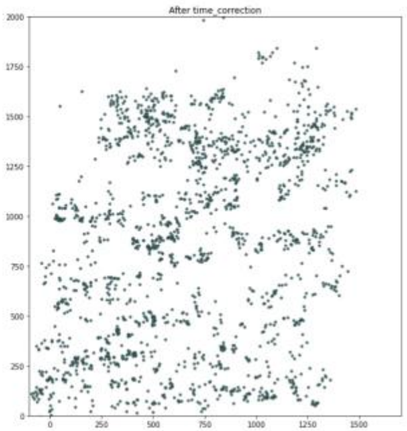
timestamp	x	y
33	1	1
66	nan	nan
99	nan	nan

case 2

timestamp	x	y
33	1	1
33	1	1
66	3	5

timestamp	x	y
33	1	1
66	3	5





About
Mission 1 : iVT Filter

Motivation
Problem Definition
Mission 2: Line allocation
Mission 3: Metric

Mission 1: iVT Filter

Time_correction ()

1. 중복 바로 뒤 timestamp - 중복값들 중 첫 timestamp > 40

: 단순 중복이 아닌 경우

1) 중복 바로 뒤 x,y값이 결측값이 아닌 경우

: 중복값에 대하여 그 자리에 x, y, timestamp를 선형으로 fill-in

2) 중복 바로 뒤 x,y값이 결측값인 경우

: timestamp는 선형으로 fill-in, x,y는 결측값 처리

→ 결측값으로 처리된 부분은 gap-fill-in에서 추가로 더 처리될 예정

2. 중복 바로 뒤 timestamp - 중복값들 중 첫 timestamp <= 40

: 중복값을 그냥 없애줌

case 1-1)

timestamp	x	y
33	1	1
33	1	1
99	3	5

case 1-2)

timestamp	x	y
33	1	1
33	1	1
99	nan	nan

case 2

timestamp	x	y
33	1	1
33	1	1
66	3	5

About
Mission 1 : iVT Filter

Motivation
Problem Definition
Mission 2: Line allocation
Mission 3: Metric

Mission 1: iVT Filter

Time_correction ()

1. 중복 바로 뒤 timestamp - 중복값들 중 첫 timestamp > 40

: 단순 중복이 아닌 경우

1) 중복 바로 뒤 x,y값이 결측값이 아닌 경우

: 중복값에 대하여 그 자리에 x, y, timestamp를 선형으로 fill-in

2) 중복 바로 뒤 x,y값이 결측값인 경우

: timestamp는 선형으로 fill-in, x,y는 결측값 처리

→ 결측값으로 처리된 부분은 gap-fill-in에서 추가로 더 처리될 예정

2. 중복 바로 뒤 timestamp - 중복값들 중 첫 timestamp <= 40

: 중복값을 그냥 없애줌

timestamp	x	y
33	1	1
66	2	3
99	3	5

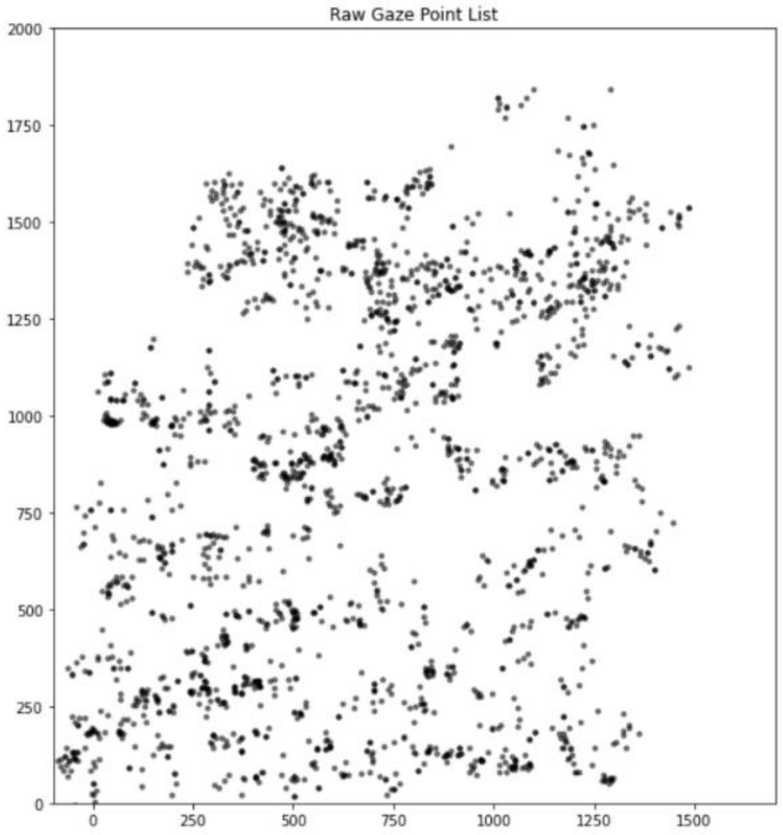
timestamp	x	y
33	1	1
66	nan	nan
99	nan	nan

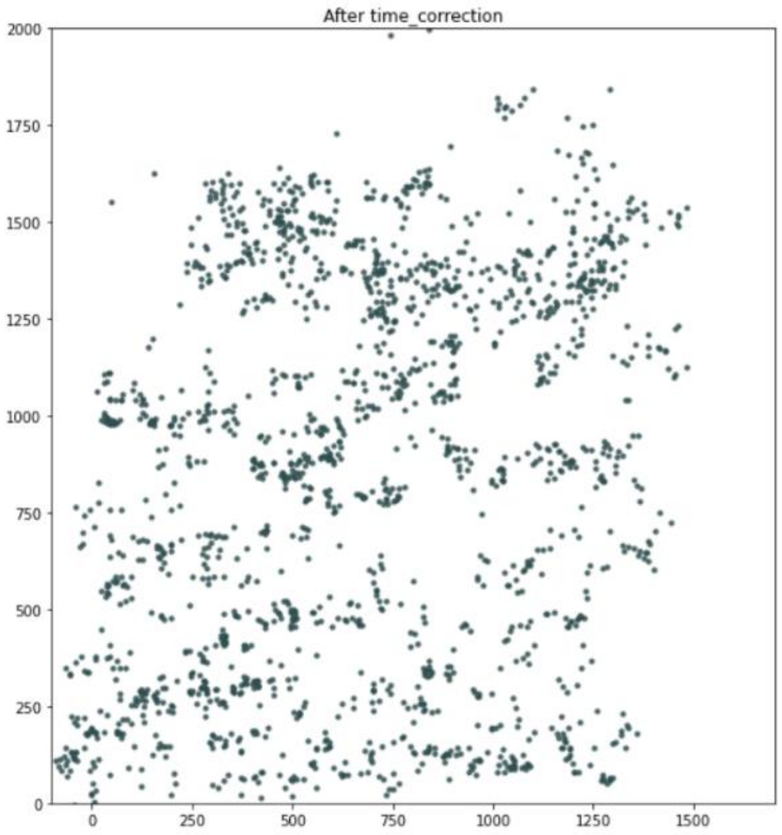
timestamp	x	y
33	1	1
66	3	5

About
Mission 1 : iVT Filter

Motivation
Problem Definition
Mission 2: Line allocation
Mission 3: Metric

After Time correction





About
Mission 1 : iVT Filter

Motivation
Problem Definition
Mission 2: Line allocation
Mission 3: Metric

Mission 1: iVT Filter

Gap_fill_in ()

목적

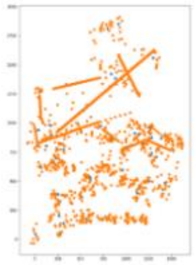
fixation의 중간에 data loss가 발생해 이를 valid data로 채워지지 않는다면 두 fixation이라고 알고리즘이 분류하게 될 것. 이를 fixation으로 묶어주기 위함

방법

1. gap의 기간 < max_gap_length
즉 1,2개씩만 결측이 연속으로 존재하는 경우 그냥 gap 자체를 drop

2. max_gap_length < gap의 기간 < max_fill_length
(1) x 좌표 차이 < max_x_distance
y 좌표 차이 < max_y_distance
: timestamp에 비례하여 선형으로 보간
(2) otherwise
: timestamp를 연속적으로 만들기

3. gap의 기간 > max_fill_length
: timestamp를 연속적으로 만들기



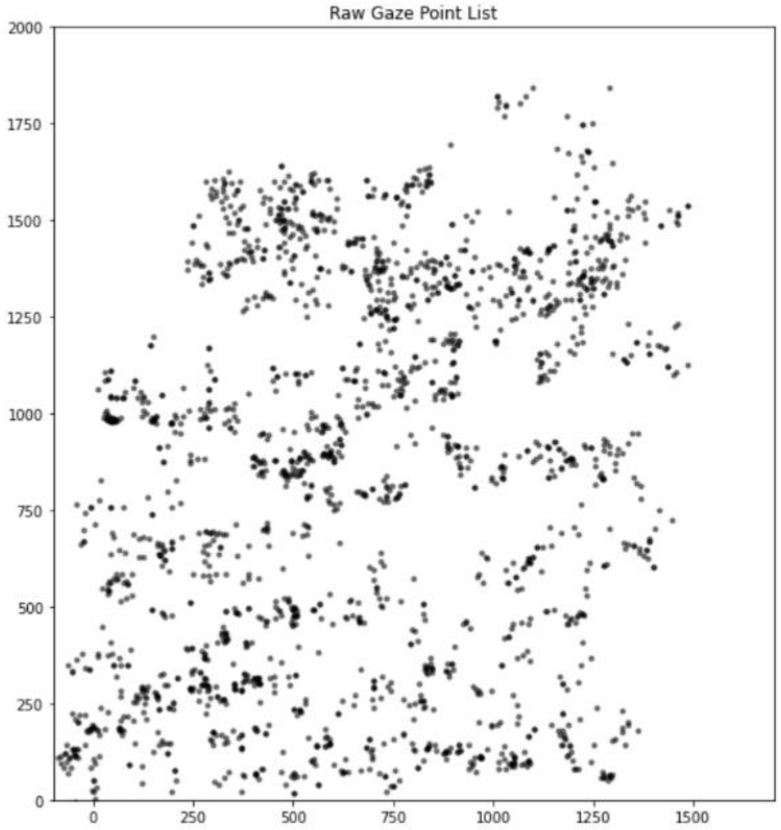
max_gap_length

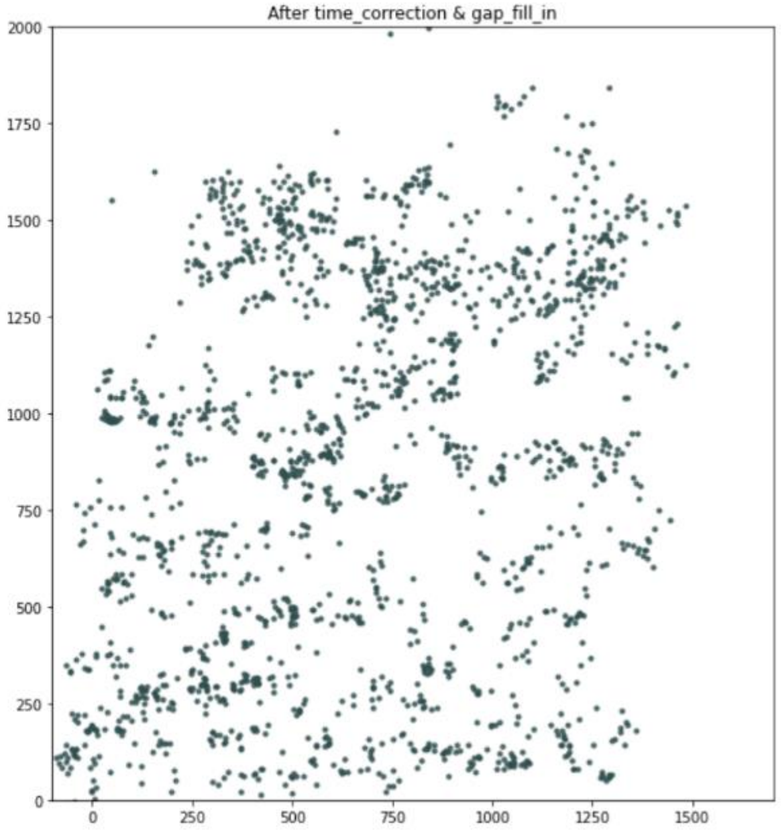
max_fill_length

About
Mission 1 : iVT Filter

Motivation
Problem Definition
Mission 2: Line allocation
Mission 3: Metric

After Time correction and Gap fill in





About
Mission 1 : iVT Filter

Motivation
Problem Definition
Mission 2: Line allocation
Mission 3: Metric

Mission 1: iVT Filter

Noise_reduction ()

목적

시선의 x축 y축 좌표를 보정하여, 속력이 튀는 값을 잘 잡기 위해서 하는 것

방법

1. 이동평균법

default로 사용하던 것 n개의 [x,y] ⇒ [(sum(xi)/n , (sum(yi)/n)]

2. 지수평활법

이동평균의 값 현재 값에 줄이 옮겨질 경우 x좌표가 크게 변화하기 때문에 alpha 값을 크게 선정

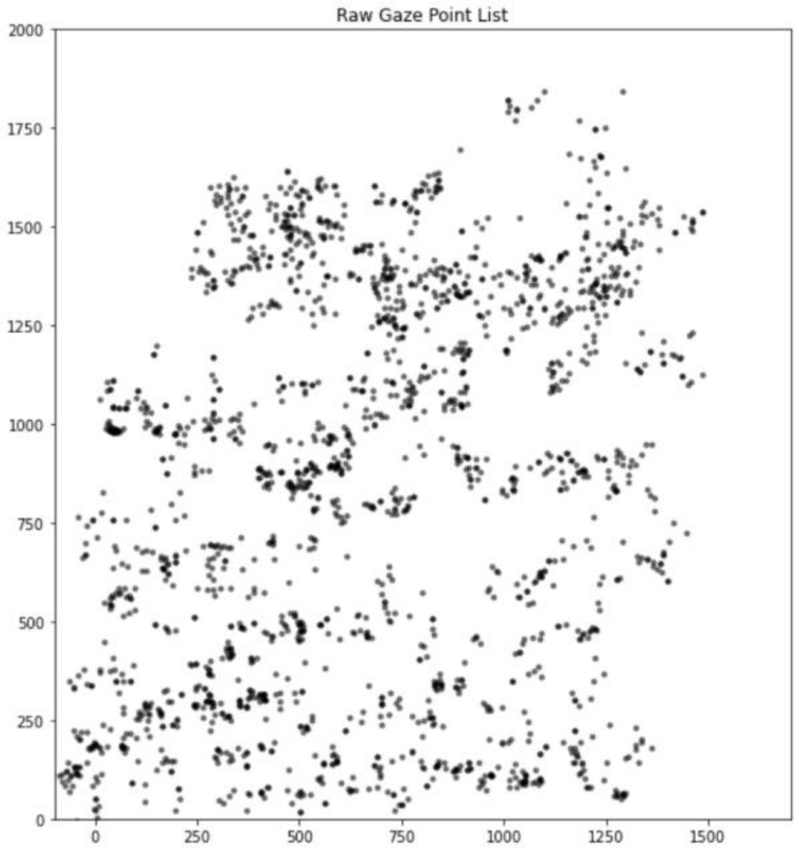
(+) 3. 칼만필터링

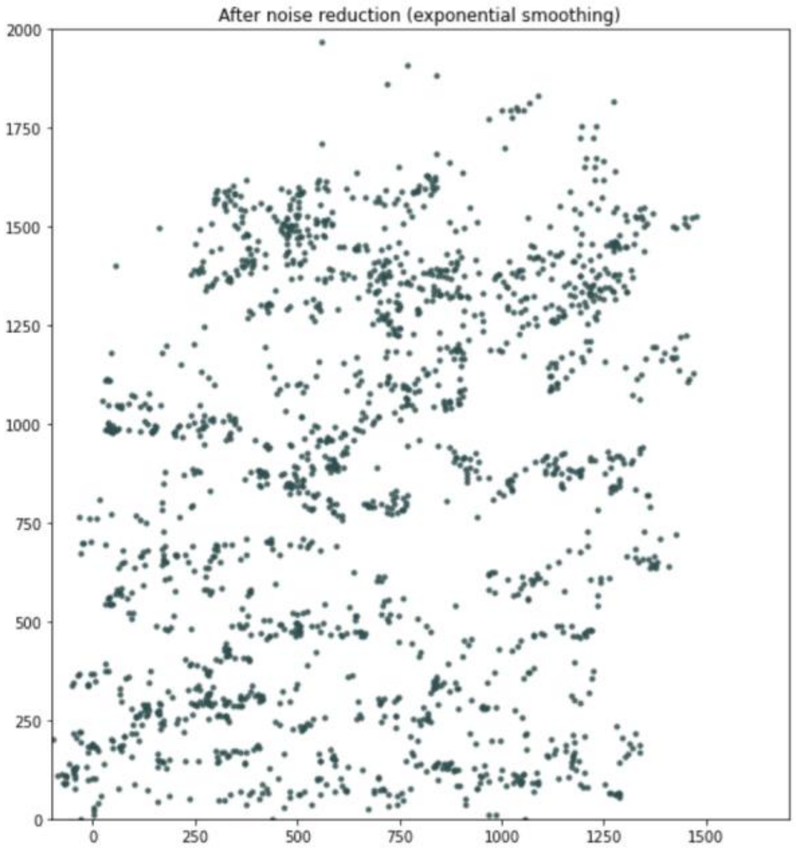
About

Mission 1 : iVT Filter

Motivation
Problem Definition
Mission 2: Line allocation
Mission 3: Metric

After Noise Reduction

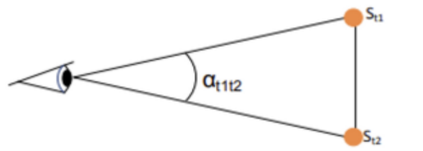




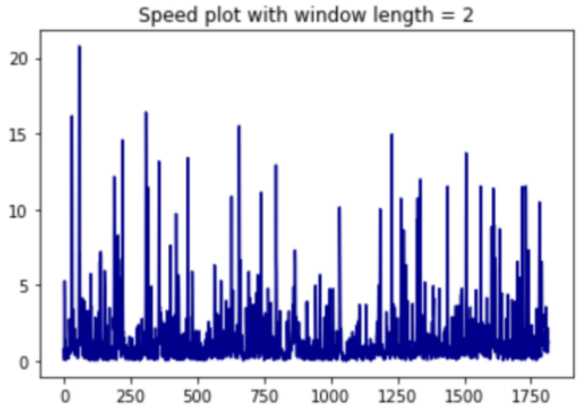
About
Mission 1 : iVT Filter

Motivation
Problem Definition
Mission 2: Line allocation
Mission 3: Metric

Mission 1: iVT Filter



Velocity_calculator ()



목적

각 gaze point 간의 속도를 계산하여 일정 속도 이상의 gaze point 들을 하나의 fixation으로 볼 것인지 결정

방법

window length=2 를 사용하여 두 gaze point 간의 x, y 좌표의 거리를 그 timestamp의 간격으로 나누어 줌.

- 각각의 gaze point마다 speed를 할당해 주는데, 두 gaze point의 속도는 앞 시점의 gaze point에 할당이 됨. 예를 들어 A (1, 1, 33), B (1, 2, 66)의 gaze point가 있다면 둘 사이의 속도인 1/33이 A (1, 1, 33) 에 할당이 됨
- window legth-1 개 만큼 결측값이 생기는데 이는 window length-1개를 처음과 끝에 결측값 인접값으로 할당함.(ex. 여기서는 1개이므로 마지막에 마지막 값 한 개 더 넣어줌)

참고: window length>2 이면, window length 개수의 점을 하나의 속도로 계산 (마지막 점, 첫번째 점의 거리와 시간 차로 구함)

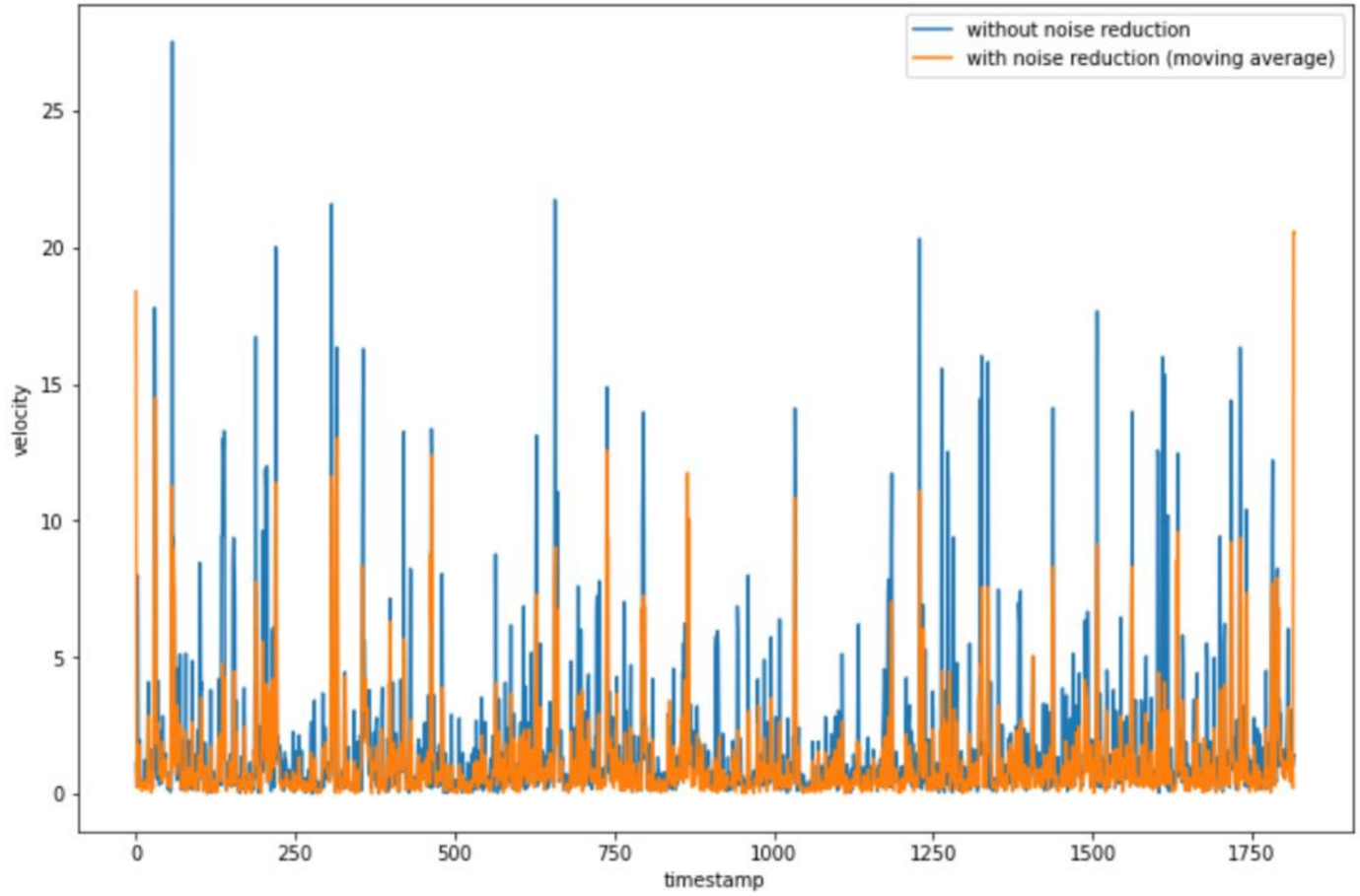
About

Mission 1 : iVT Filter

Motivation
Problem Definition
Mission 2: Line allocation
Mission 3: Metric

Calculate velocity

w/ noise reduction vs w/o noise reduction



About
Mission 1 : iVT Filter

Motivation
Problem Definition
Mission 2: Line allocation
Mission 3: Metric

Mission 1: iVT Filter

iVT_Filter ()

1. classification saccade or fixation

방법

if **gaze_point.speed > velocity_threshold : saccade**

이전 시점의 속도가 매우 크기 때문에 시선이 도약한 것이라고 판단

if **gaze_point.speed < velocity_threshold : fixation**

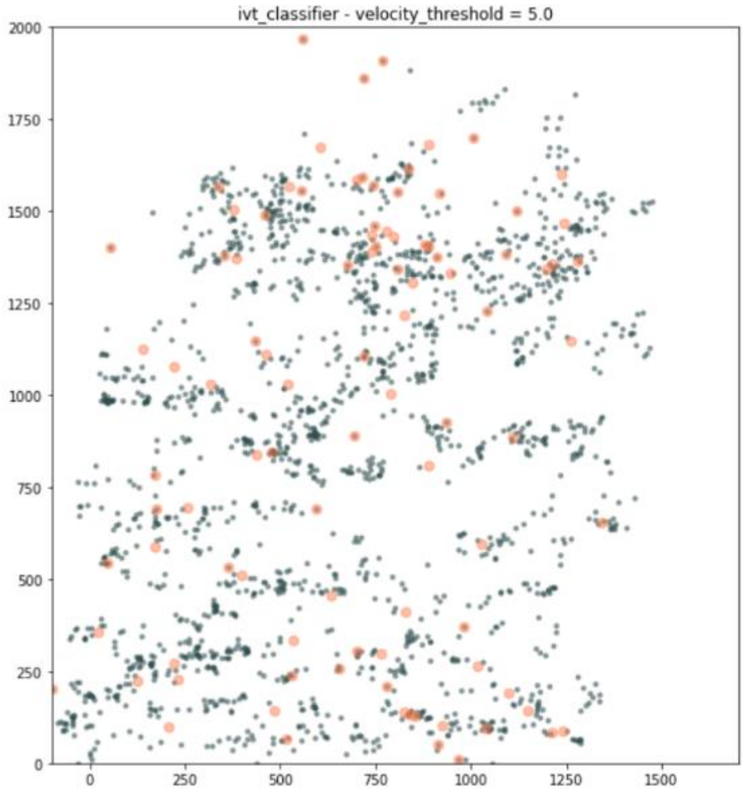
이전 시점의 속도가 작기 때문에 시선이 한 지점에 머물러 있는 것이라고 판단

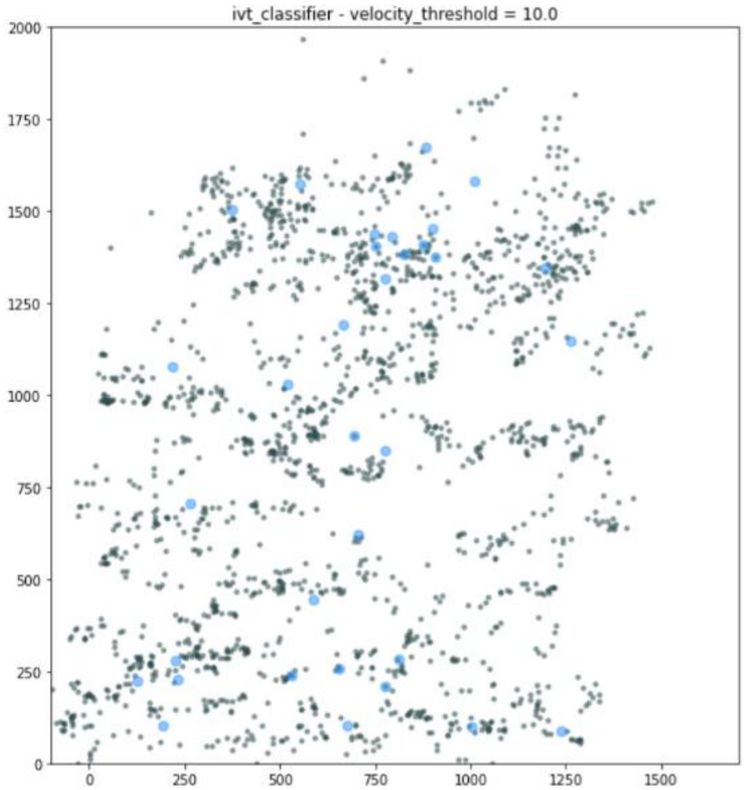
About

Mission 1 : iVT Filter

Motivation
Problem Definition
Mission 2: Line allocation
Mission 3: Metric

Comparasion-velocity threshold





About
Mission 1 : iVT Filter

Motivation
Problem Definition
Mission 2: Line allocation
Mission 3: Metric

Mission 1: iVT Filter

iVT_Filter ()

2. merge adjacent fixation

목적

fixation의 위치 정보만을 고려하여 인접한 위치의 fixation을 하나의 fixation으로 merge 한다면
Inline regression을 catch 하지 못하기 때문에 **위치 정보 + 시간 정보를 모두 고려**해야 한다.

방법

1. 맨 앞의 fixation과 그 다음에 존재하는 fixation들의 euclidean distance를 구하여 timestamp 순으로 나열한다.
2. timestmap 순으로 distance를 확인하며 merge_threshold_distance 보다 작다면 같은 fixation으로 묶는다
3. merge_threshold_distance 보다 큰 distance를 가진 fixation이 생긴다면 그 fixation부터는 merge하지 않는다.
4. 그 fixation 을 기준으로 1~3을 모든 fixation에 대해 확인할 때까지 반복한다.

About
Mission 1 : iVT Filter

Motivation
Problem Definition
Mission 2: Line allocation
Mission 3: Metric

Mission 1: iVT Filter

iVT_Filter ()

2. merge adjacent fixation

Euclidean Distance Matrix에서

	Fix 1	Fix 2	Fix 3	Fix 4	Fix 5	Fix 6	Fix 7
Fix 1	0	50	90	120	80	200	250
Fix 2	50	0	30	70	200	120	350
Fix 3	90	30	0	35	120	140	180
Fix 4	120	70	35	0	130	150	150
Fix 5	80	200	120	130	0	20	60
Fix 6	200	120	140	150	20	0	35
Fix 7	250	350	180	150	60	35	0

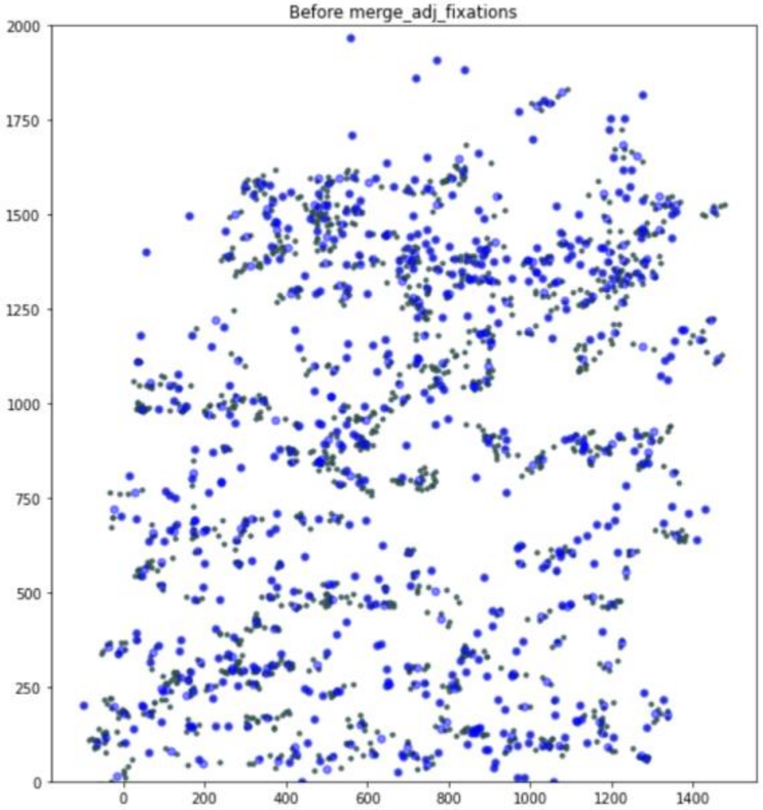
- Fix 1을 기준으로 먼저 살펴보면 Fix 3까지는 모두 100 이하이므로 같은 fixation group으로 merge
- Fix 4와의 거리가 120으로 threshold보다 크기 때문에 Fix 4부터 다시 merge할 fixation 찾기
- Fix 4는 Fix 5와의 거리가 130으로 threshold보다 크기 때문에 Fix 4는 혼자 group
- Fix 5는 그 다음의 모든 fixation group과의 거리가 모두 100보다 작기 때문에 5,6,7은 한 group으로 merge

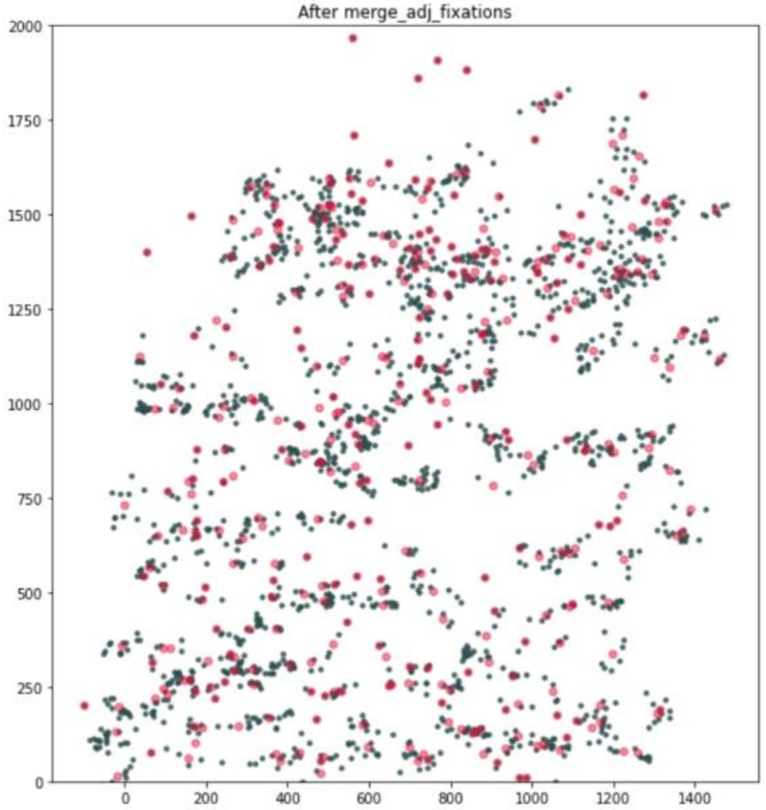
About

Mission 1 : iVT Filter

Motivation
Problem Definition
Mission 2: Line allocation
Mission 3: Metric

After Merge adj fixation





About
Mission 1 : iVT Filter

Motivation
Problem Definition
Mission 2: Line allocation
Mission 3: Metric

Mission 1: iVT Filter

iVT_Filter ()

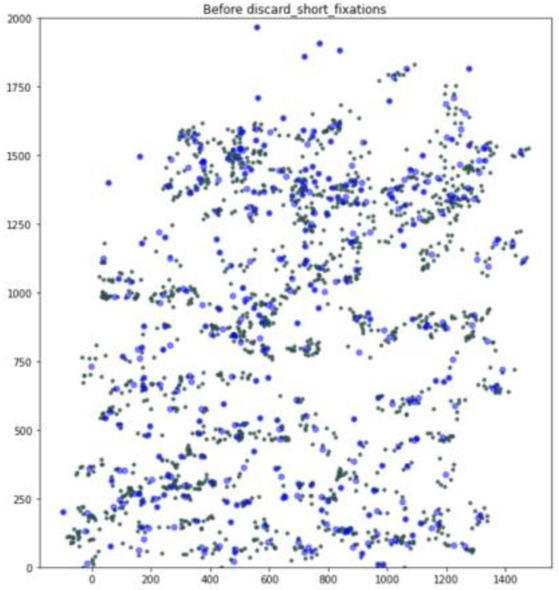
3. discard short fixation

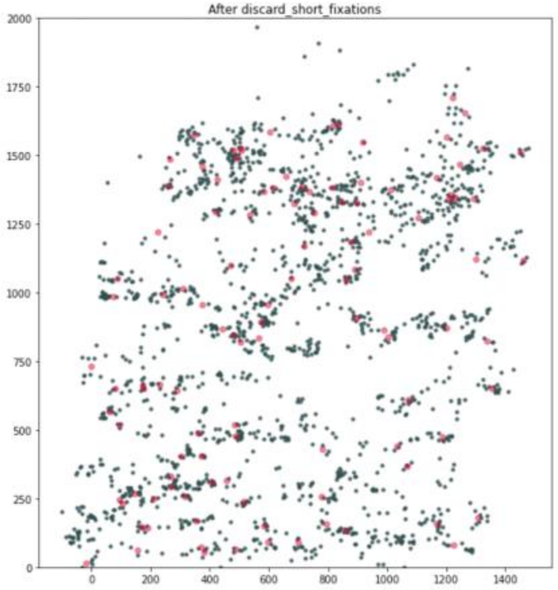
목적

불필요한 fixation 제거

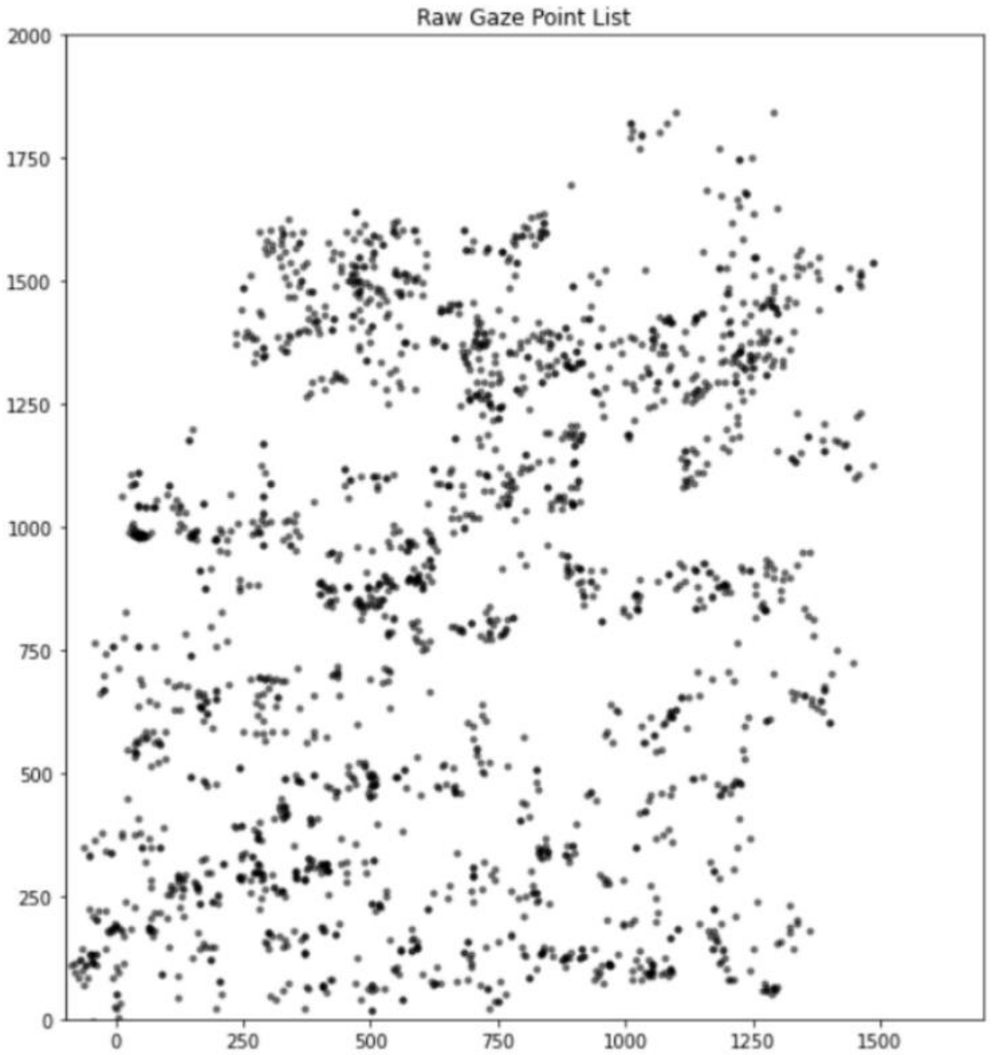
방법

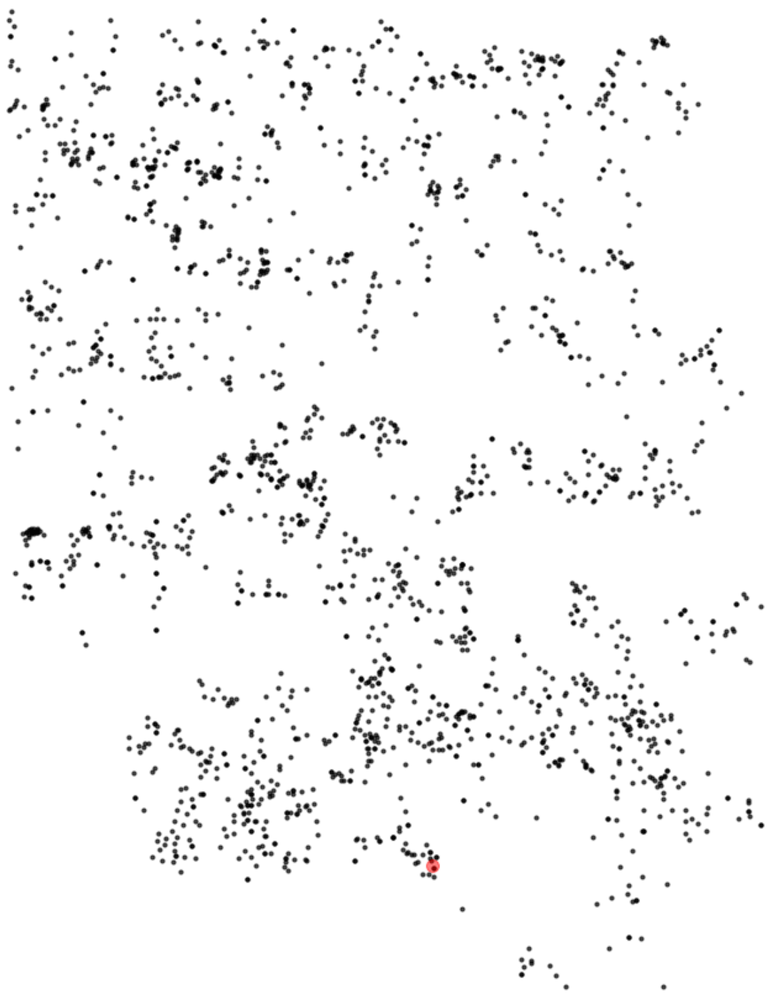
(1) fixation에 포함된 점이 1개인 fixation 제거
(2) time duration이 60 이하인 fixation 제거





Mission 1: iVT Filter 결과





About
Mission 2 : Line allocation

Motivation
Problem Definition
Mission 1: iVT Filter
Mission 3: Metric

Mission 2: Line allocation

보안 상의 이유로
공개할 수 없는 내용입니다.

About
Mission 2 : Line allocation

Motivation
Problem Definition
Mission 1: iVT Filter
Mission 3: Metric

Mission 2: Line allocation

Fixation 종류

보안 상의 이유로
공개할 수 없는 내용입니다.

About
Mission 2 : Line allocation

Motivation
Problem Definition
Mission 1: iVT Filter
Mission 3: Metric

Mission 2: Line allocation

Flow: 전처리

보안 상의 이유로
공개할 수 없는 내용입니다.

About
Mission 2 : Line allocation

Motivation
Problem Definition
Mission 1: iVT Filter
Mission 3: Metric

Mission 2: Line allocation

Flow: 전처리

보안 상의 이유로
공개할 수 없는 내용입니다.

About
Mission 2 : Line allocation

Motivation
Problem Definition
Mission 1: iVT Filter
Mission 3: Metric

Mission 2: Line allocation

Flow: 전처리

보안 상의 이유로
공개할 수 없는 내용입니다.

About
Mission 2 : Line allocation

Motivation
Problem Definition
Mission 1: iVT Filter
Mission 3: Metric

Mission 2: Line allocation

Flow: 전처리

보안 상의 이유로
공개할 수 없는 내용입니다.

About

Mission 2 : Line allocation

Motivation
Problem Definition
Mission 1: iVT Filter
Mission 3: Metric

Flow : 전처리 예시

보안 상의 이유로
공개할 수 없는 내용입니다.

About
Mission 2 : Line allocation

Motivation
Problem Definition
Mission 1: iVT Filter
Mission 3: Metric

Flow : 전처리 예시

보안 상의 이유로
공개할 수 없는 내용입니다.

About
Mission 2 : Line allocation

Motivation
Problem Definition
Mission 1: iVT Filter
Mission 3: Metric

Flow : 전처리 예시

보안 상의 이유로
공개할 수 없는 내용입니다.

About
Mission 2 : Line allocation

Motivation
Problem Definition
Mission 1: iVT Filter
Mission 3: Metric

Flow : 전처리 예시

보안 상의 이유로
공개할 수 없는 내용입니다.

About
Mission 2 : Line allocation

Motivation
Problem Definition
Mission 1: iVT Filter
Mission 3: Metric

Flow : 전처리 예시

보안 상의 이유로
공개할 수 없는 내용입니다.

About
Mission 2 : Line allocation

Motivation
Problem Definition
Mission 1: iVT Filter
Mission 3: Metric

Mission 2: Line allocation

보안 상의 이유로
공개할 수 없는 내용입니다.

About
Mission 2 : Line allocation

Motivation
Problem Definition
Mission 1: iVT Filter
Mission 3: Metric

Mission 2: Line allocation

보안 상의 이유로
공개할 수 없는 내용입니다.

About
Mission 2 : Line allocation

Motivation
Problem Definition
Mission 1: iVT Filter
Mission 3: Metric

Mission 2: Line allocation

보안 상의 이유로
공개할 수 없는 내용입니다.

About
Mission 2 : Line allocation

Motivation
Problem Definition
Mission 1: iVT Filter
Mission 3: Metric

Mission 2: Line allocation

보안 상의 이유로
공개할 수 없는 내용입니다.

About
Mission 2 : Line allocation

Motivation
Problem Definition
Mission 1: iVT Filter
Mission 3: Metric

Mission 2: Line allocation

X좌표 배정 1. Word Padding

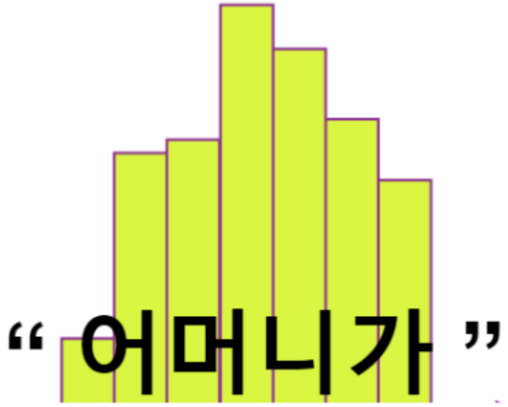
띄어쓰기 **gap의 절반**을 기준으로 padding

About
Mission 2 : Line allocation

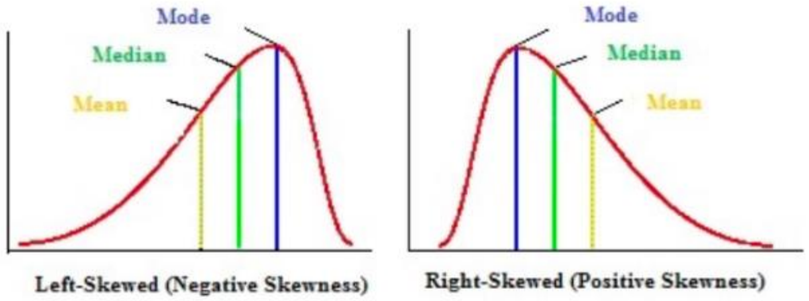
Motivation
Problem Definition
Mission 1: iVT Filter
Mission 3: Metric

Mission 2: Line allocation

X좌표 배정 2. Check the Word Distribution - Median



의미 부분 위주로 skewed일 가능성 높음



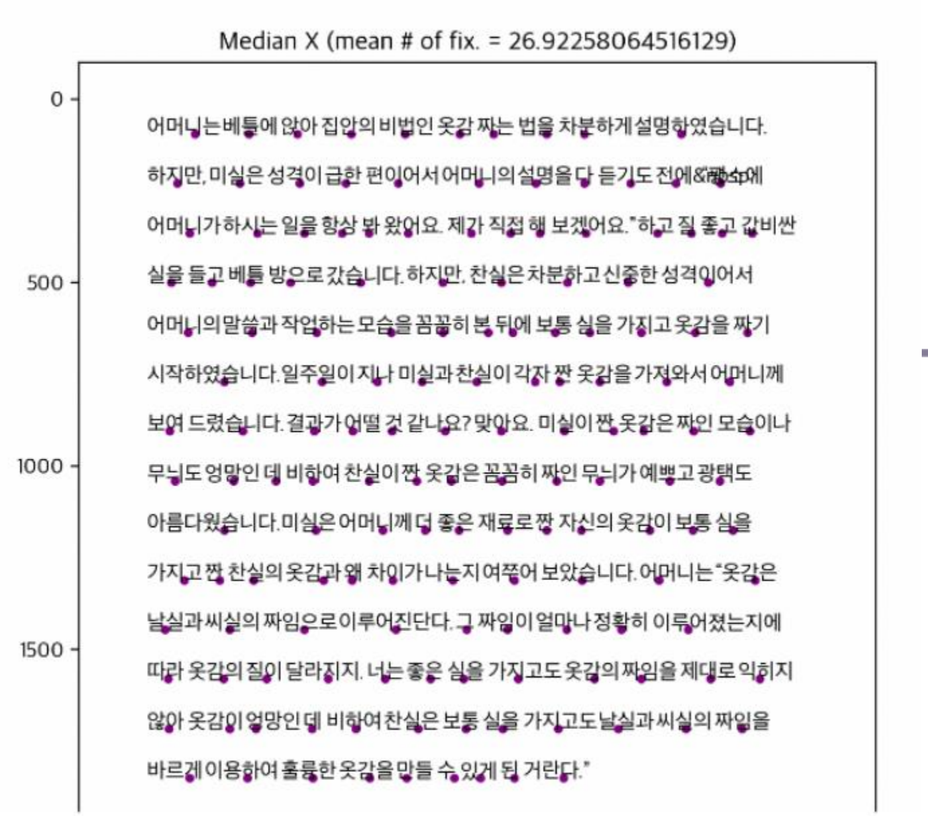
Median 사용하자!

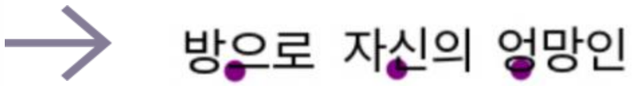
About
Mission 2 : Line allocation

Motivation
Problem Definition
Mission 1: iVT Filter
Mission 3: Metric

Mission 2: Line allocation

X좌표 배정 2. Check the Word Distribution - Median





About

Mission 2 : Line allocation

Motivation
Problem Definition
Mission 1: iVT Filter
Mission 3: Metric

Mission 2: Line allocation

보안 상의 이유로
공개할 수 없는 내용입니다.

Mission 2: Line allocation 결과

보안 상의 이유로
공개할 수 없는 내용입니다.

Mission 2: Line allocation 결과

보안 상의 이유로
공개할 수 없는 내용입니다.

Mission 2: Line allocation 결과

보안 상의 이유로
공개할 수 없는 내용입니다.

DSL
X
Visual Camp

기업 연계 프로젝트

 # Part 2 - Metric

Mission 1&2 Summary

보안 상의 이유로
공개할 수 없는 내용입니다.

Mission 1&2 Summary

보안 상의 이유로
공개할 수 없는 내용입니다.

About
Mission 3 : Metric

Motivation
Problem Definition
Mission 1: iVT Filter
Mission 2: Line allocation

Mission 3: Metric

보안 상의 이유로
공개할 수 없는 내용입니다.

About
Mission 3 : Metric

Motivation
Problem Definition
Mission 1: iVT Filter
Mission 2: Line allocation

Mission 3: Metric

보안 상의 이유로
공개할 수 없는 내용입니다.

About
Mission 3 : Metric

Motivation
Problem Definition
Mission 1: iVT Filter
Mission 2: Line allocation

Mission 3: Metric

선행 연구 정리

A Meta-Analysis of Eye-Tracking Studies on Text Processing in Children with Reading Disabilities (2018)

: Jung A Kim, Se Jin Oh, Eunjung Choi, Young Tae Kim, Jee Eun Sung

[결론]
- **읽기장애 아동 집단**이 일반아동 집단에 비해 유의하게 **더 많은 고정 수와 더 긴 고정 지속시간**을 보인 것으로 나타났다.
- **읽기장애 아동 집단**이 일반아동 집단에 비해 유의하게 **더 많은 도약 수**를 보인 것으로 나타났다.
- **읽기장애 아동 집단**이 일반아동 집단에 비해 유의하게 **더 많은 회귀 수**를 보인 것으로 나타났다.

Reading Comprehension and Reading Processing of School-Aged Children with Specific Language Impairment Using Eye Tracker (2018)

: Si-Nae Kang, Dongsun Yim

[결론]
- **단순언어장애 아동**은 일반아동에 비해 이야기를 읽는 데 유의하게 더 오랜 시간이 필요하였으며, **평균 시선고정 시간과 총 시선 고정 시간이 유의하게 긴 것**으로 나타났다. 단순언어장애 아동의 **총 시선고정 횟수 또한 일반아동에 비해 유의하게 많은 것**으로 나타났다.
- 읽기 시간과 시선고정 횟수, 총 시선고정 시간과 읽기 이해력 사이에 부적 상관관계가 나타났다. 즉, **읽기 이해력에서 높은 점수를 받은 아동일수록 읽기 시간 및 시선고정 시간이 짧고 시선고정 횟수가 적게 나타나는 경향**을 보였으며. 이를 통해 높은 읽기 이해력을 보이는 아동은 효율적인 읽기를 할 가능성이 높다는 것을 확인 가능하였다.

About
Mission 3 : Metric

Motivation
Problem Definition
Mission 1: iVT Filter
Mission 2: Line allocation

Mission 3: Metric

Metric 1: 지문별 평균 단어 이해 속도

	time	word	speed
age			
8	70108.447368	84.0	834.624373
9	67207.250000	103.0	652.497573
10	71710.977273	133.0	539.180280
11	72354.802817	155.0	466.805179
12	74512.645161	160.0	465.704032
13	82119.137500	156.0	526.404728

} 저학년 평균 : 675.434

} 고학년 평균 : 486.305

About
Mission 3 : Metric

Motivation
Problem Definition
Mission 1: iVT Filter
Mission 2: Line allocation

Mission 3: Metric

Metric 2: Duration mean

age	time	word	duration mean
8	70108.447368	84.0	376.793135
9	67207.250000	103.0	366.749903
10	71710.977273	133.0	338.091936
11	72354.802817	155.0	330.500073
12	74512.645161	160.0	323.700475
13	82119.137500	156.0	322.298704

} 저학년 평균 : 360.545

} 고학년 평균 326.000

About
Mission 3 : Metric

Motivation
Problem Definition
Mission 1: iVT Filter
Mission 2: Line allocation

Mission 3: Metric

Metric 3: 시폭

Forward Reading으로 태깅된 fixation 간의 거리 비교

	forward 시폭 평균
age	
8	154.302075
9	134.572971
10	145.237283
11	116.377128
12	141.200795
13	123.880116

} 저학년 평균 : 144.704

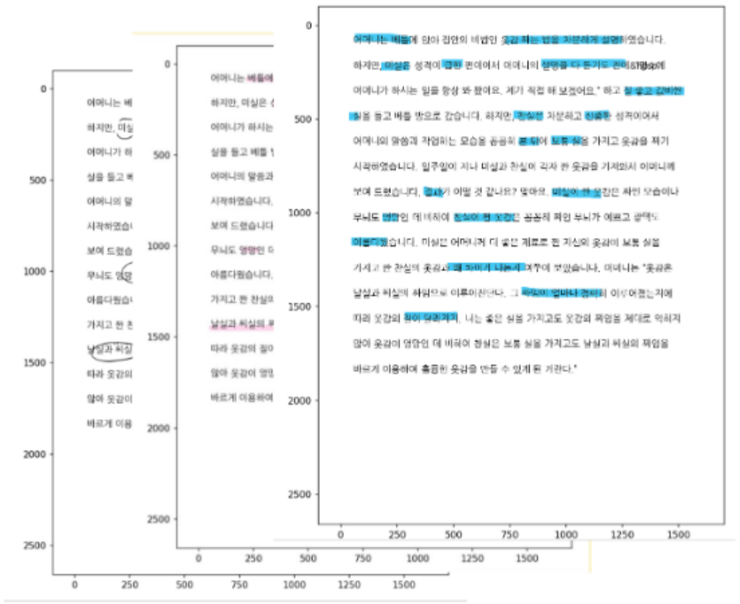
} 고학년 평균: 127.152

About
Mission 3 : Metric

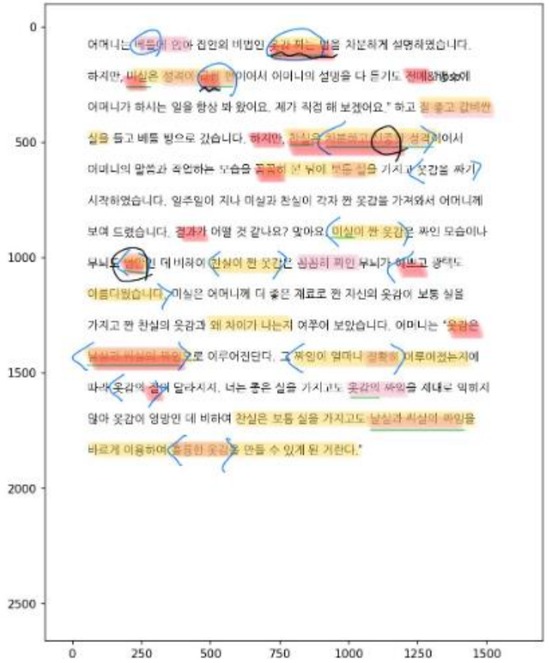
Motivation
Problem Definition
Mission 1: iVT Filter
Mission 2: Line allocation

Mission 3: Metric

Metric 4: Major Word 이용(new!)



- 핵심 단어 안에 fixation이 얼마나 포함되어 있는지를 파악하자!
- 지문을 파악하기 위한 핵심 단어, Major Word 추출
(ex. 중요도: 급한 > 편이어서)



(한 사람 빼고) 전부 동일하게 체크한 부분 = Major

About
Mission 3 : Metric

Motivation
Problem Definition
Mission 1: iVT Filter
Mission 2: Line allocation

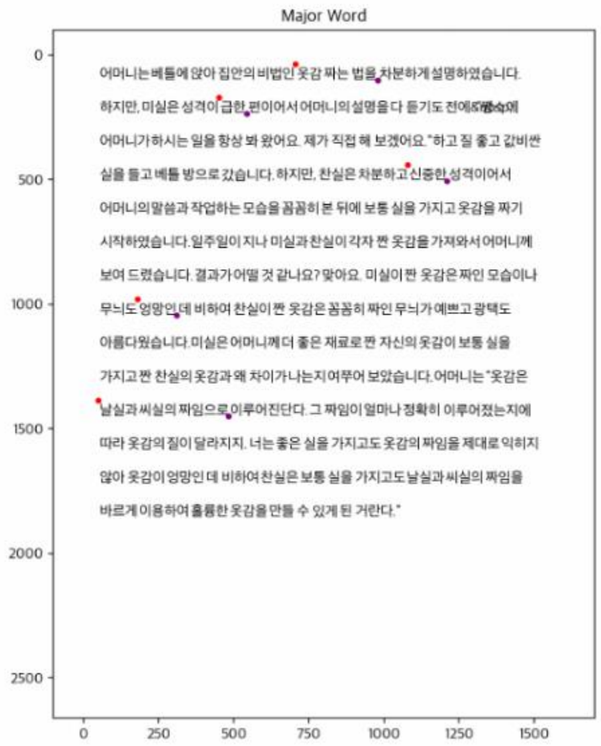
Mission 3: Metric

Metric 4: Major Word 이용

Major Word

- 옷감 짜는 법을
- 급한
- 신중한
- 엉망인
- 날실과 씨실의 짜임

+ Padding →



About
Mission 3 : Metric

Motivation
Problem Definition
Mission 1: iVT Filter
Mission 2: Line allocation

Mission 3: Metric

Metric 4: Major Word 이용

Major Word

- 옷감 짜는 법을
- 급한
- 신중한
- 엉망인
- 날실과 씨실의 짜임

IN major words	IN major words (5point)
[0, 0, 1, 0, 0]	1
[1, 1, 1, 1, 1]	5
[1, 1, 0, 0, 0]	2
[0, 0, 1, 0, 1]	2
[0, 0, 0, 1, 0]	1
[0, 0, 0, 0, 1]	1
[0, 0, 0, 0, 0]	0
[0, 0, 1, 0, 1]	2
[0, 0, 1, 0, 1]	2
[0, 0, 0, 1, 1]	2
[0, 0, 1, 1, 0]	2
[0, 0, 0, 0, 0]	0
[0, 0, 1, 0, 1]	2
[0, 0, 1, 0, 1]	2
[0, 0, 1, 1, 1]	3
[0, 0, 1, 0, 1]	2
[0, 0, 0, 0, 1]	1
[0, 1, 0, 0, 1]	2
[0, 1, 1, 1, 1]	4

각 단어 영역에 fixation이 존재하는가
: 0(없다), 1(있다)로 표현

평균 1.9/5점
(1, 2점이 합해서 48명으로 다수)
5점 : 1명
0점 : 6명

About
Mission 3 : Metric

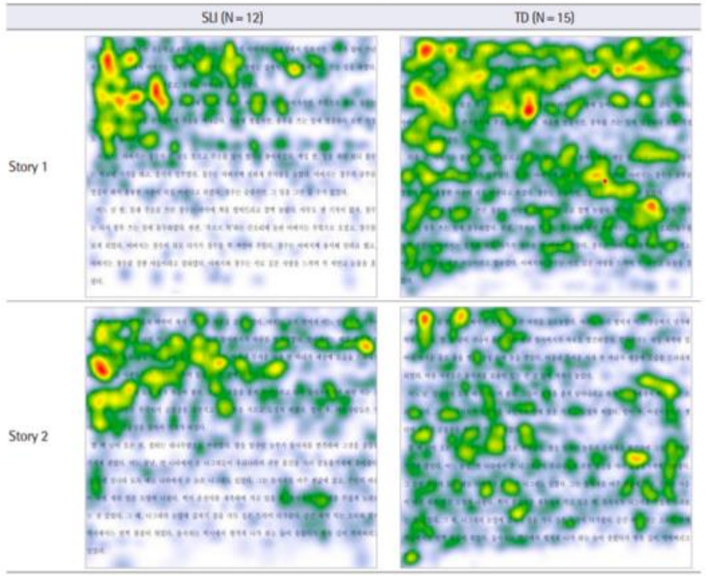
Motivation
Problem Definition
Mission 1: iVT Filter
Mission 2: Line allocation

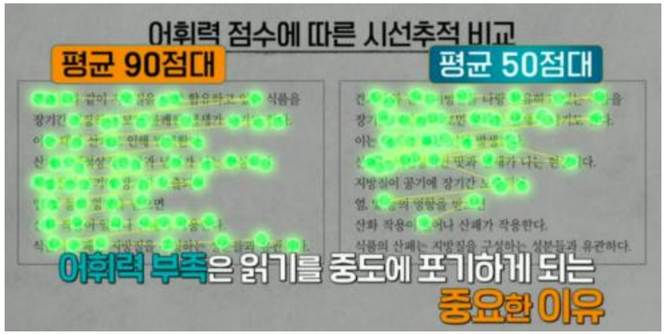
Mission 3: Metric

Metric 5: Fixation의 분포 이용 (In progress)

문해력에 문제가 있는 집단의 경우, 아래 그림과 같이 후반부에는 시선이 많이 머물지 않는 것을 볼 수 있음

- 이야기를 읽는 과정에서 텍스트의 전반부에서 많은 시간을 할애하여 정보처리를 한 반면, 후반부에서는 읽기 집중도가 떨어졌다고 해석 가능
- Fixation의 분산, Duration의 감소폭 등을 새로운 Metric의 기준으로 세워볼 수 있을 것이라 판단, 진행 중





Q&A

THANK YOU