

Wyniki B+R

CV.RAP12 Raport z badania porównawczego wygenerowanej mapy uwagi wizualnej z wynikami badania okulistycznego

© 2025 CinemaVision. All rights reserved.

Spis treści

1.	Uzasadnienie przeprowadzenia badania porównawczego wygenerowanej mapy uwagi wizualnej z wynikami badania okulistycznego na tym samym materiale filmowym	2
2.	Metryka badania	3
2.1.	Materiał badawczy	3
2.2.	Cele badania	3
2.3.	Narzędzia badawcze	3
2.3.1.	Narzędzie Movie Diagnostics Saliency Map	3
2.3.2.	Stanowisko i narzędzie ET Diagnostic do badania okulistycznego	4
2.4.	Badanie okulistyczne	4
2.4.1.	Dobór grupy badanych	4
2.4.2.	Scenariusz przeprowadzenia badania okulistycznego	4
2.4.3.	Rezultaty badania okulistycznego	4
2.5.	Generowanie map uwagi	5
2.5.1.	Agregacja wyników generowania mapy i badania okulistycznego	5
2.6.	Porównanie wygenerowanych map uwagi z wynikami okulistyki per ujęcie	7
2.7.	Wskaźniki korelacji pomiędzy wygenerowaną mapą, a wynikami badania okulistycznego	39
3.	Podsumowanie i wnioski	40
3.1.	Przetestowanie skuteczności algorytmów budujących mapę uwagi – wnioski	40
3.2.	Analiza przypisywania wag algorytmom w tworzeniu mapy uwagi - wnioski	40
3.3.	Porównanie mapy uwagi z wynikami badania okulistycznego – wnioski	41
4.	Ocena przydatności metody komputerowego generowania mapy uwagi dla materiału filmowego do badania preferencji widzów	41

1. Uzasadnienie przeprowadzenia badania porównawczego wygenerowanej mapy uwagi wizualnej z wynikami badania okulograficznego na tym samym materiale filmowym

Kluczowe definicje:

Atencja wizualna – zdolność selekcji pewnego zbioru informacji z otoczenia.

Mapa uwagi (mapa punktów skupienia wzroku, mapa punktów fiksacji wzroku) – to wizualizacja, która reprezentuje uwagę wizualną obrazu. Mapa uwagi jest generowana jako dwuwymiarowa bitmapa w skali szarości. W prezentowanym badaniu mapa uwagi jest wizualizowana w postaci heatmapy, gdzie kolor niebieski oznacza niską uwagę wizualną, a kolor czerwony wysoką. Mapa różnicuje obszary obrazu pod względem zdolności przyciągania wzroku i koncentracji uwagi. Wygenerowana mapa uwagi wizualnej pozwala na rozpoznanie potencjalnie problematycznych pod względem przyciągania uwagi obszarów materiału filmowego na podstawie analizy otrzymanych wyników.

Badanie okulograficzne – śledzenie i rejestracja ruchów gałek ocznych. Badanie to polega na śledzeniu pierwszego postrzegania określonych elementów za pomocą specjalnej kamery, rejestrującej najmniejsze ruchy gałki ocznej. Badanie pomaga zweryfikować to, w jaki sposób ludzie postrzegają obiekty na obrazie, który śledzą. Dostarcza informacji gdzie w danym momencie badany patrzy, na których elementach się skupia, a które całkowicie pomija. Ma to na celu sprawdzenie czy odbiorca patrząc na określony obraz widzi te elementy, na których interesariuszom zależy.

Badanie porównawcze wygenerowanych map uwagi wizualnej z mapą uzyskaną za pomocą badania okulograficznego przeprowadzono w celu sprawdzenia skuteczności zaproponowanej metody komputerowego tworzenia map obszarów kadru w materiale filmowym, który będzie przyciągał wzrok potencjalnego widza. Skuteczność metody będzie rozumiana jest jako współczynnik korelacji pomiędzy wygenerowaną mapą uwagi wizualnej, a mapą przedstawiającą dane okulograficzne badanej grupy.

Zarówno mapa uwagi wizualnej jak i mapa przedstawiające wyniki badania okulograficznego mają postać czarno-białej bitmapy (pliku graficznego z pojedynczej klatki filmu). W przypadku danych z okulografu obszary białe oznaczają miejsca, w które spojrzała się osoba, a czarne – miejsca, w które nie spojrzała. W przypadku mapy uwagi wizualnej, obszary białe oznaczają miejsca, w które ktoś spojrzy z pewnym prawdopodobieństwem, a obszary czarne – miejsca, w które nikt się nie spojrzy.

2. Metryka badania

2.1. Materiał badawczy

	Showreel Isyrius Długość: • 00:01:42:08 • 2457 klatek Wideo: • Kodex: H264 – MPEG-4 AVC • Rozdzielczość HD 1920: 800 • Klatkaż 24 klatki na sekundę • Pixel Aspect Ratio 1,0 • Zdekodowany format: Planar 4:2:0 YUV • Bit rate 1524 kb/s Audio: • Kodex: MPEG AAC Audio • Kanał: Stereo • Częstotliwość próbkowania 48000 Hz
---	--

Jako materiał badawczy został użyty materiał wideo Showreel Isyrius o długości 1 m 42 s (2457 klatek), w rozdzielczości fullHD. Materiał zawiera różnorodne wizualnie krótkie ujęcia, które pozwalają sprawdzić uniwersalność proponowanej metody.

Showreel Isyrius jest materiałem marketingowym firmy Isyrius zajmującej się kompleksową produkcją materiałów audiowizualnych „od pomysłu do realizacji” – filmów fabularnych, dokumentalnych, reklamowych, korporacyjnych, teledysków, etiud i innych. Materiał stanowi przegląd dotychczasowych realizacji wykonanych w różnych okresach. Cechuje go szybki montaż oraz bogactwo ujęć.

2.2. Cele badania

- Przetestowanie skuteczności algorytmów budujących mapę uwagi:
 - Przetestowanie wyglądu mapy uwagi w zależności od złożoności sceny i sekwencji wydarzeń w kadrze.
 - Poszukiwanie trendów i aproksymacja oczekiwań w zależności od zawartości kadru.
 - Ocena wartości analitycznej uzyskanych wyników dla poszczególnych algorytmów metody.
- Analiza koncepcji przypisywania różnych wag algorytmom w tworzeniu wynikowej mapy uwagi:
 - Oszacowanie istotności i trafności poszczególnych algorytmów składowych metody w tworzeniu wynikowej mapy uwagi.
 - Analiza wpływu zastosowanych wag w zależności od charakteru analizowanego ujęcia.
- Porównanie mapy uwagi z wynikami badania okulograficznego:
 - Ocena trafności wygenerowanych map uwagi na podstawie mapy uwagi uzyskanej w wyniku badania okulograficznego na tym samym materiale badawczym (filmowym).

2.3. Narzędzia badawcze

2.3.1. Narzędzie Movie Diagnostics Saliency Map

Movie Diagnostic Saliency Map to aplikacja, która umożliwia komputerowe wyliczenie i wygenerowanie 8-bitowego rozkładu mapy uwagi wizualnej dla zadanego obrazu/kadru lub klatki materiału filmowego.

Prawdopodobieństwo wystąpienia obszarów fiksacji wzroku dla pojedynczej klatki filmowej jest wyliczane na podstawie:

- analizy rozkładu jasności i kontrastu w obrazie,

- analizy wielkości i rozłożenia obiektów w kadrze,
- detekcji wystąpienia twarzy w obrazie,
- detekcji ruchu obiektów w płaszczyźnie ekranu.

Wygenerowana mapa uwagi jest konwertowana na heatmapę, która ma szeroką rozpiętość kolorystyczną – od niebieskiego, przez żółty, do czerwonego. Prezentacja mapy uwagi polega na wygenerowaniu pliku filmowego (graficznego) z wyróżnionymi obszarami uwagi na tle materiału źródłowego.

2.3.2. Stanowisko i narzędzie ET Diagnostic do badania okulograficznego

Badanie przeprowadzono za pomocą urządzenia rejestrującego Tobii EyeX i autorskiego oprogramowania ET Diagnostic składającego się z modułu rejestracji miejsc skupienia wzroku badanego oraz modułu wizualizującego położenie miejsc skupienia wzroku na tle badanego filmu (uwaga wzroku). Wizualizacje mogą prezentować wyniki dla poszczególnych badanych oraz wyniki zbiorcze dla wszystkich..

Stanowisko badawcze zostało zorganizowane w kameralnej sali konferencyjnej. Sala była klimatyzowana, izolowana akustycznie od dźwięków zakłócających komfortowy odbiór filmu. Badany i moderator siedzieli obok siebie przy dużym stole konferencyjnym, na którym był rozłożony sprzęt pomiarowy i było wydzielone wygodne miejsce do instalacji urządzenia do eyetrackingu Tobii.

2.4. Badanie okulograficzne

2.4.1. Dobór grupy badanych

Grupę badawczą w badaniu stanowiły osoby niezwiązane ze środowiskiem filmowym, nieposiadające wysokich kwalifikacji w zakresie technik realizacyjnych i montażowych oraz budowania ekspresyjności przekazu. Przyjęto założenie, że „świeże spojrzenie” jest dobrym kryterium identyfikacyjnym dla badanych i rokuje utrzymanie koncentracji w trakcie oglądania całego badanego materiału filmowego.

Badanie przeprowadzono na grupie reprezentatywnej 18 osób w przedziale wieku 23-42. Grupa badawcza składała się z 10 kobiet i 8 mężczyzn.

2.4.2. Scenariusz przeprowadzenia badania okulograficznego

1. Przywitanie badanego, wprowadzenie w kontekst badania i przedstawienie celu badania i oczekiwanych czynności ze strony badanego oraz harmonogramu wydarzeń w trakcie badania.
2. Zapoznanie badanego z eyetrackerem Tobii EyeX, wyjaśnienie istoty jego działania i zaspokojenie ciekawości związanej z celowością jego użycia.
3. Skalibrowanie urządzenia Tobii EyeX do parametrów wzroku badanego. Test kontrolny skanowania wzrokiem całości ekranu emisyjnego.
4. Emisja materiału badawczego i obserwacja przebiegu badania.
5. Zakończenie badania, sprawdzenie poziomu skuteczności zebrania danych okulograficznych. W przypadku poziomu braku danych powyżej 8% lub czasu pomiędzy zbieranymi danymi powyżej 300 ms powtórzenie badania.
6. Podziękowanie badanemu za udział w badaniu, odpowiedzi na dodatkowe pytania.

2.4.3. Rezultaty badania okulograficznego

Badanie okulograficzne rejestrowało położenie wzroku badanego (punkt fiksacji) z częstotliwością 60 Hz. Start zbierania danych był zsynchronizowany programowo z momentem rozpoczęcia emisji badanego filmu. Dane były agregowane w identyfikowalnym, unikalnym dla konkretnego badanego pliku csv zawierającym następujące informacje:

- moment czasowy od startu badania,
- koordynaty x,y położenia punktu fiksacji.

Wszystkie badania zawierały wystarczającą ilość danych do zakwalifikowania ich do przeprowadzenia analizy zbiorczej.

Wizualizacja wyników grupowych z badania okulograficznego posłużyła do zbudowania zbiorczej mapy uwagi wzroku dla grupy badawczej nałożonej na film. Mapa stanowi materiał porównawczy dla wygenerowanych za pomocą aplikacji CinemaVision Saliency Map map uwagi. Zgodność obu materiałów posłużyła do analizy trafności metody komputerowego generowania mapy uwagi. Wysoka trafność będzie potwierdzeniem zasadności użycia generatora mapy uwagi wizualnej w badaniach preferencji widzów dotyczących uwagi materiałów filmowych.

2.5. Generowanie map uwagi

W badaniu wygenerowano następujące mapy uwagi:

- składową mapę uwagi na podstawie analizy rozkładu jasności i kontrastu w obrazie (contrast, intensity) w skali 0-255, w wersji czarno-białej,
- składową mapę uwagi na podstawie analizy wielkości i rozłożenia obiektów w kadrze (size and distribution) w skali 0-255, w wersji czarno-białej,
- składową mapę uwagi na podstawie detekcji wystąpienia twarzy w obrazie (faces) w skali 0-255, w wersji czarno-białej,
- składową mapę uwagi na podstawie detekcji ruchu obiektów w płaszczyźnie ekranu (motion) w skali 0-255, w wersji czarno-białej.
- skumulowaną mapę uwagi na podstawie przypisania wag określonym powyżej mapom składowym w wersji kolorowej heatmapy w skali 0-255 w kolorach od niebieskiego do czerwonego. Mapa skumulowana zawierała następujące wagi składowe (3-2-2-3):
 - 30% (contrast,intensity),
 - 20% (size and distribution),
 - 20% (faces),
 - 30% (motion),
- skumulowaną mapę uwagi na podstawie przypisania wag określonym powyżej mapom składowym w wersji kolorowej heatmapy w skali 0-255 w kolorach od niebieskiego do czerwonego. Mapa skumulowana zawierała następujące wagi składowe (1-2-2-5):
 - 10% (contrast,intensity),
 - 20% (size and distribution),
 - 20%(faces),
 - 50%(motion),
- skumulowaną mapę uwagi na podstawie grupowych wyników badania okulograficznego (eyetracking map) w wersji kolorowej heatmapy w skali 0-255 w kolorach od czerwonego (niska kumulacja uwagi) do białego (wysoka kumulacja uwagi).

Mapy uwagi zostały wygenerowane w postaci filmów zsynchronizowanych z badanym materiałem. W przypadku mapy (faces) warstwa detekcyjna została nałożona jako półprzezroczysta warstwa na oryginalny materiał badany w celu łatwiejszej identyfikacji położenia zidentyfikowanych obszarów uwagi.

2.5.1. Agregacja wyników generowania mapy i badania okulograficznego

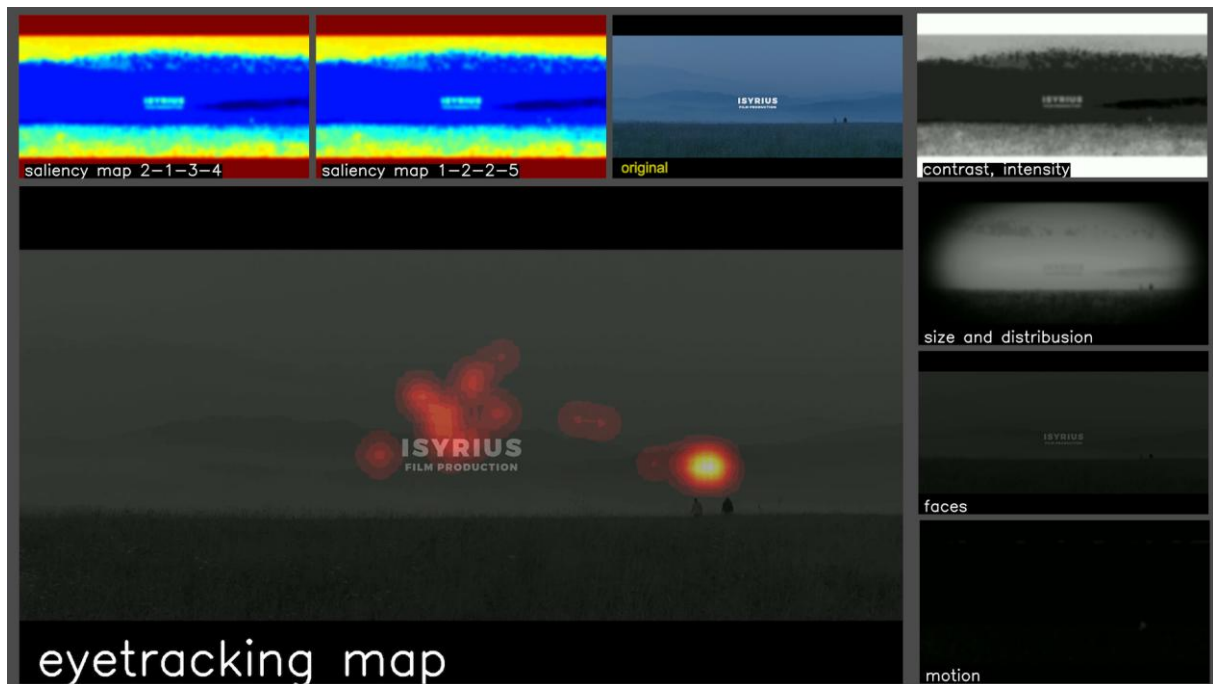
Schemat organizacji materiału wizualnego z badania wygląda następująco:

saliency map dla bilansu wag 30-20-20-30 (%)	saliency map dla bilansu wag 10-20-20-50 (%)	obraz źródłowy	saliency map wyłącznie na bazie analizy rozkładu jasności i kontrastu w obrazie (waga 100%)
mapa uwagi z badań okulograficznych			saliency map wyłącznie na bazie analizy wielkości i rozłożenia obiektów w kadrze (waga 100%)
			saliency map wyłącznie na bazie detekcji wystąpienia twarzy w obrazie (waga 100%)

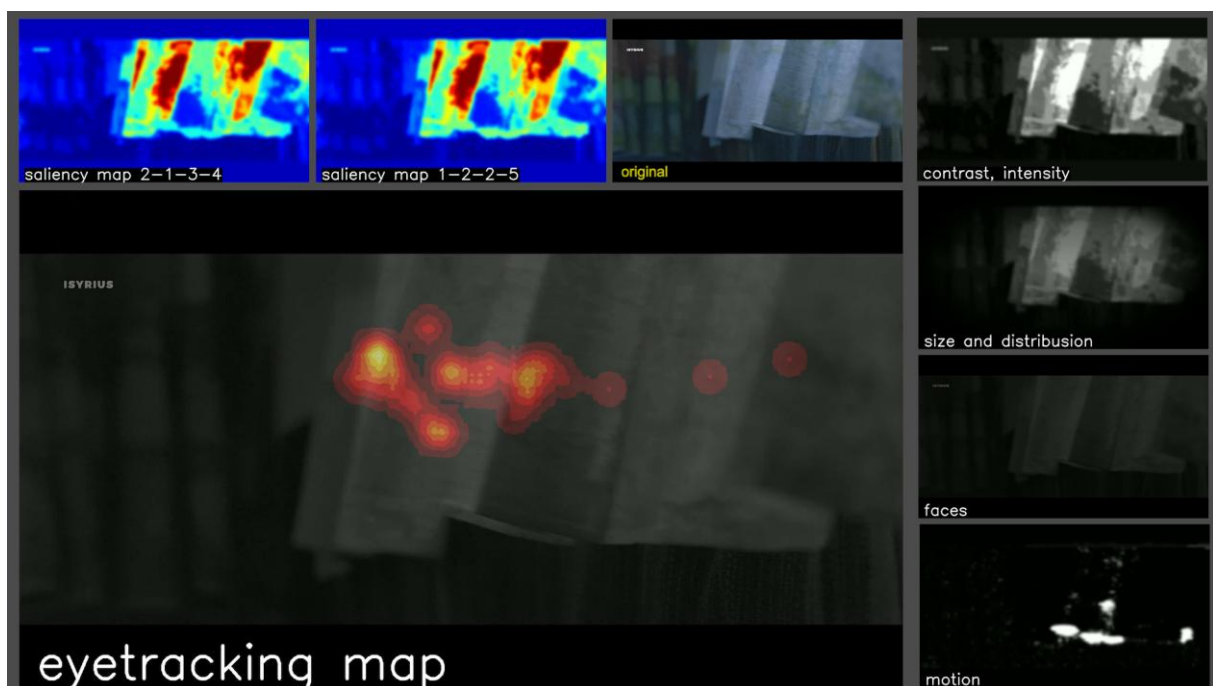


Rys. Kompozycja kadru filmu prezentującego zestawienie wyników badania.

2.6. Porównanie wygenerowanych map uwagi z wynikami okulografii per ujęcie

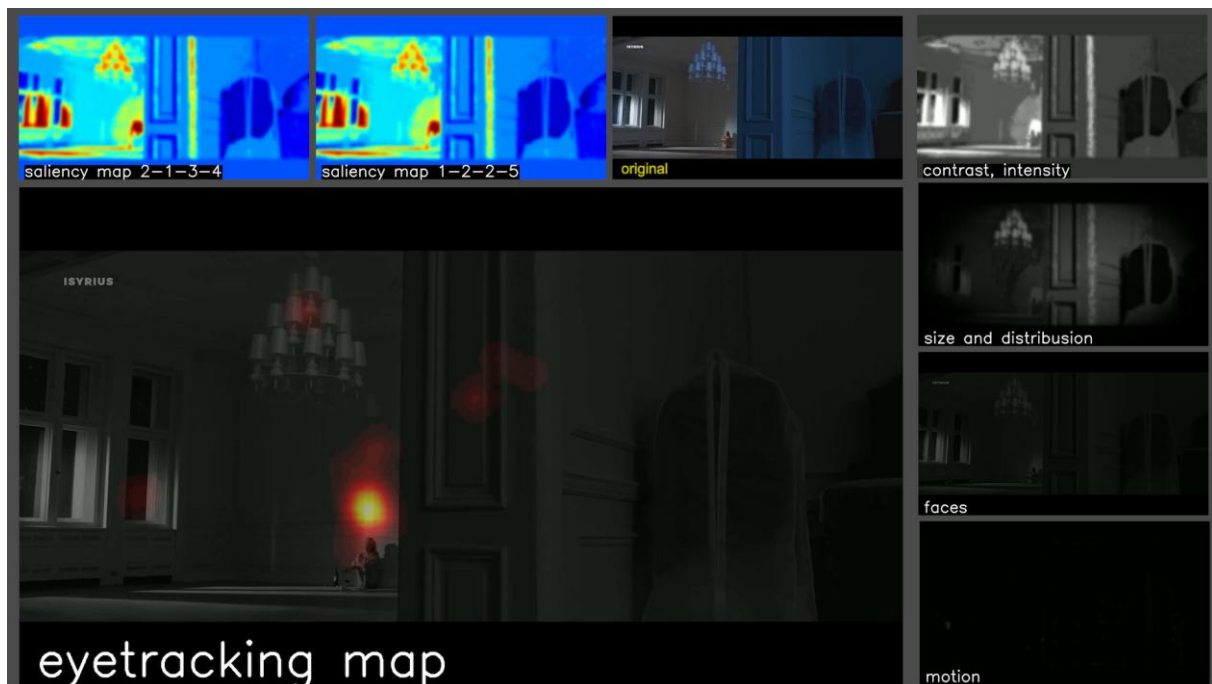
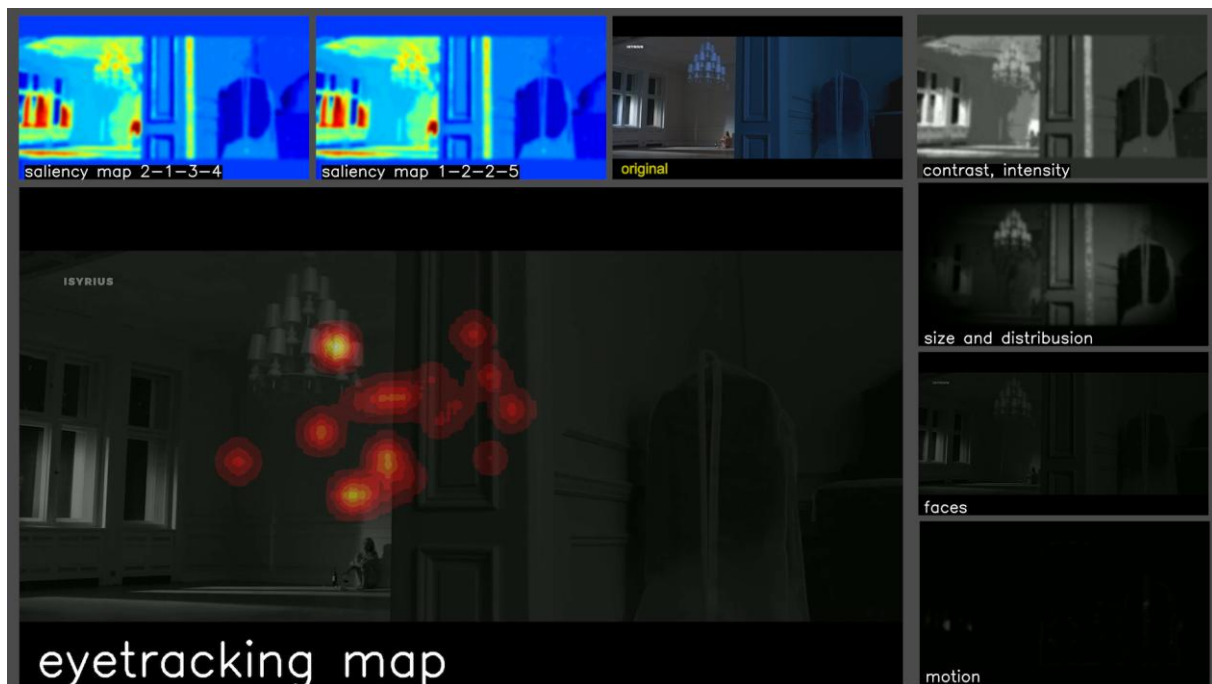


Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
0033	Czołówka, daleki pusty krajobraz, w tle idące dwie małe postacie.	Koncentracja wokół tytułu i postaci.	Rozpoznanie tytułu i postaci.	Czarne pasy (góra oraz dół) wpływają na złe wyniki mapy "contrast, intensity". Wykryty ruch ma za małą wagę.	30%

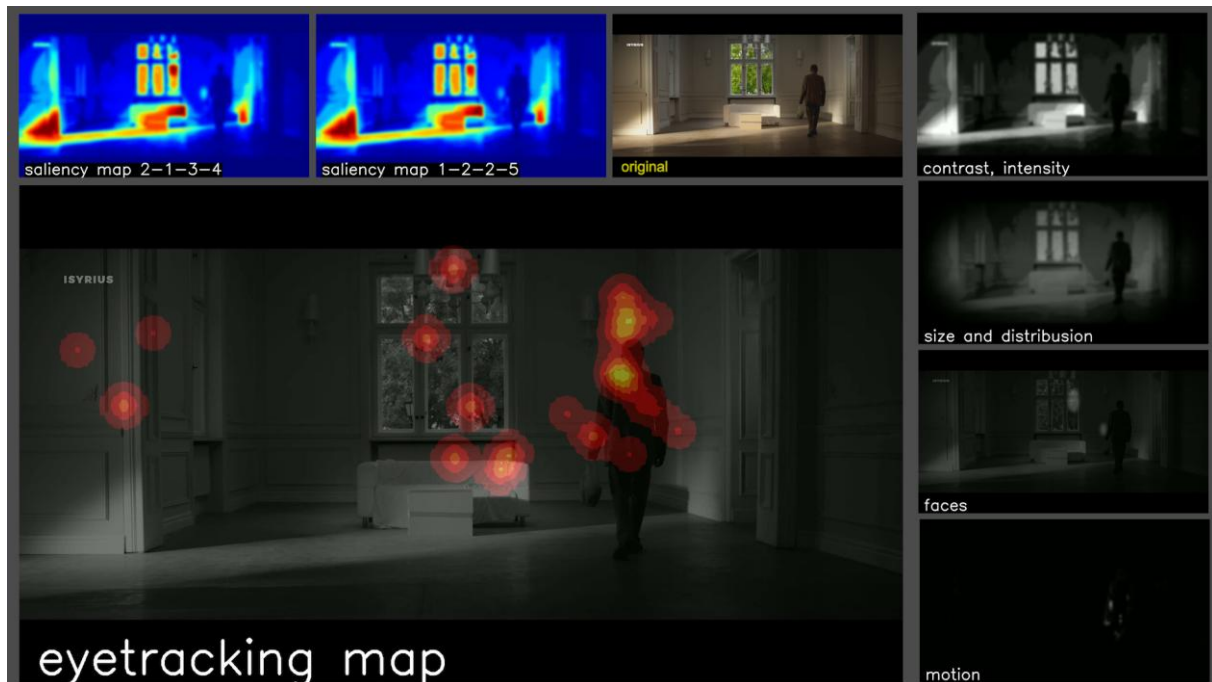


Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
0033	Czołówka, daleki pusty krajobraz, w tle idące dwie małe postacie.	Koncentracja wokół tytułu i postaci.	Rozpoznanie tytułu i postaci.	Czarne pasy (góra oraz dół) wpływają na złe wyniki mapy "contrast, intensity". Wykryty ruch ma za małą wagę.	30%

0108	Firanka poruszająca się na wietrze.	Koncentracja na ruchu firanki.	Rozpoznany ruch.	Ruch wykryto tylko na krawędziach firanki.	90%
------	-------------------------------------	--------------------------------	------------------	--	-----



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
0160 0169	Scena statyczna, półmrok, skulona mała postać na drugim planie.	Koncentracja wokół twarzy postaci. Analiza pomieszczenia.	Identyfikacja obiektów.	Sama postać pozostaje niewykryta (za to obszar w pobliżu postaci już tak). Brak rozpoznania ubrania po prawej. Żyrandol rozpoznany ze zbyt dużą wagą w stosunku do postaci.	70%



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
0243 0286	Scena statyczna, postacie w centrum kadru, półmrok.	Idąca postać, centrum kadru, badanie obiektów na scenie.	Kanapa oraz okno (centrum kadru, jasne obszary).	Nie wykryto postaci (jest odwrócona). Przypisano za dużą wagę światłu na podłodze.	50%



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
0342	Scena statyczna chata i las, nieruchoma postać z lewej strony.	Koncentracja wokół twarzy, analiza sceny.	Scena jest skomplikowana i uwaga widzów jest rozproszona, co jest odzwierciedlone przez mapę uwagi.	Brak wykrycia postaci i twarzy.	40%



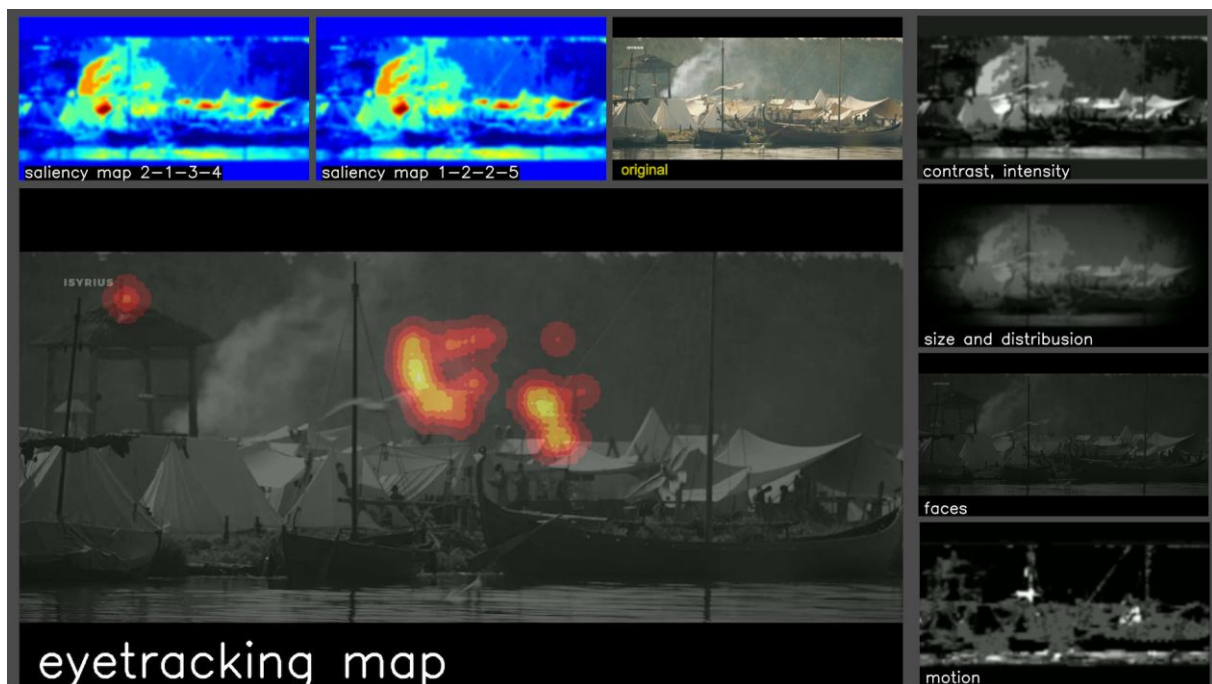
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
0366	Scena statyczna, postacie kobiece w centrum kadru, półmrok, światła punktowe.	Koncentracja wokół twarzy, później analiza obiektów na scenie.	Rozpoznana twarz na wprost, identyfikacja obiektów.	Zbyt duża waga dla całej postaci. Wykryta tylko jedna twarz.	75%



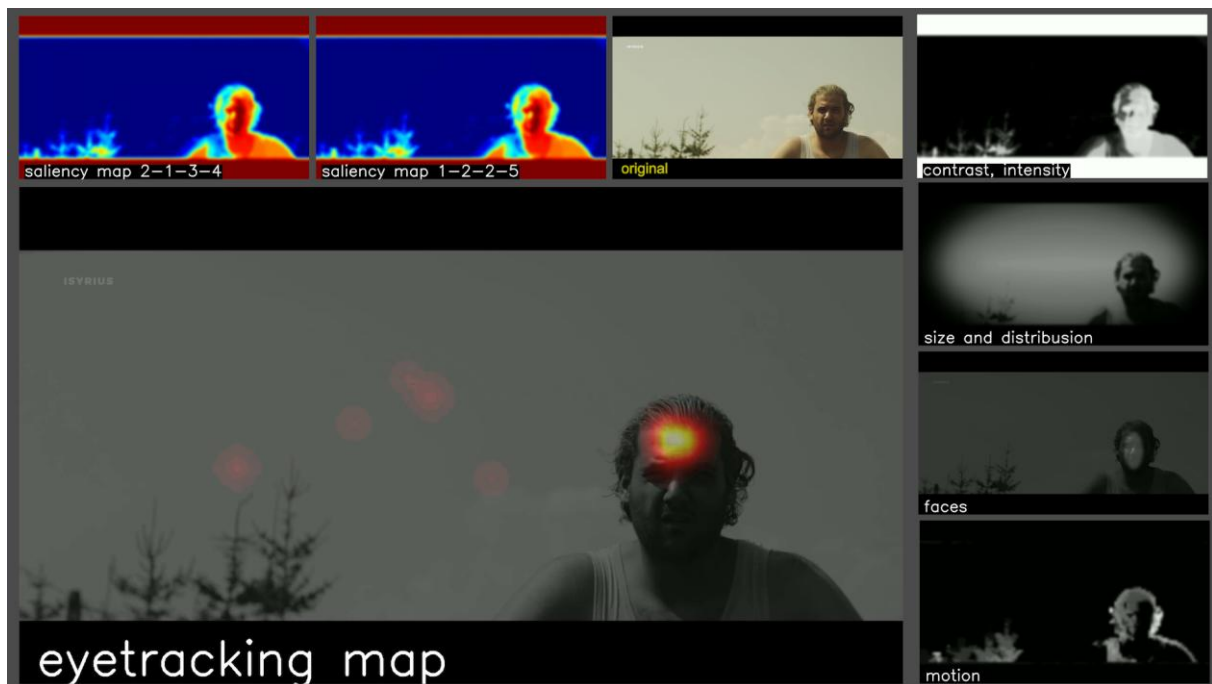
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
0430 0462	Scena statyczna, kobieta w centrum poprawia sukienkę.	Koncentracja wokół twarzy, obiekty na scenie (ubrania).	Koncentracja na środku sceny, na wyróżniających się lub jasnych obiektach.	Brak wykrycia twarzy postaci.	80%



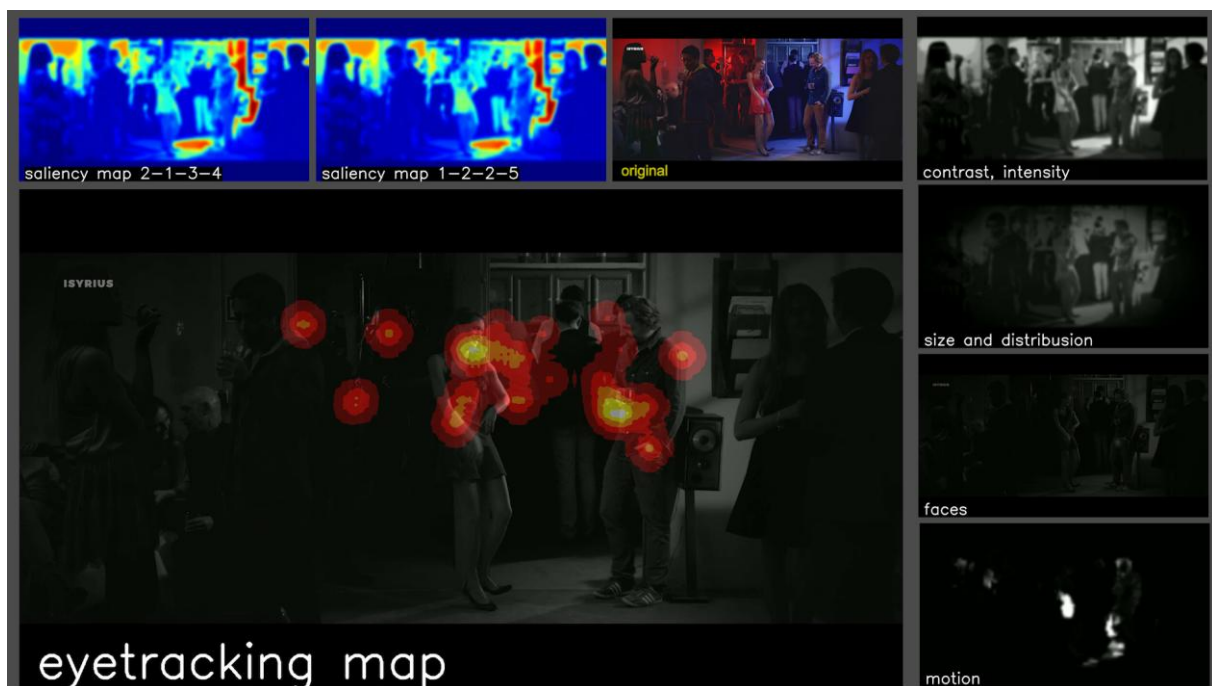
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
0513	Plan pełny statyczny, postacie w centrum kadru, dużo światła.	Koncentracja wokół twarzy, analiza sceny.	Rozpoznana twarz mężczyzny, koncentracja na pierwszym planie.	Brak wykrycia twarzy kobiety, twarz wykryta w złym miejscu, twarz powinna bardziej przyciągać uwagę niż reszta rzeczy.	60%



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
0602	Plan ogólny, plener, przelatujące ptaki na tle statków.	Koncentracja wokół lecących ptaków.	Rozpoznany ruch ptaków, łódki oraz dym przyciągają uwagę.	W końcowej mapie ruch ma za małą wagę (niknie przyćmiony przez kontrast i intensywność, które w tym przypadku nie odgrywają kluczowej roli).	70%



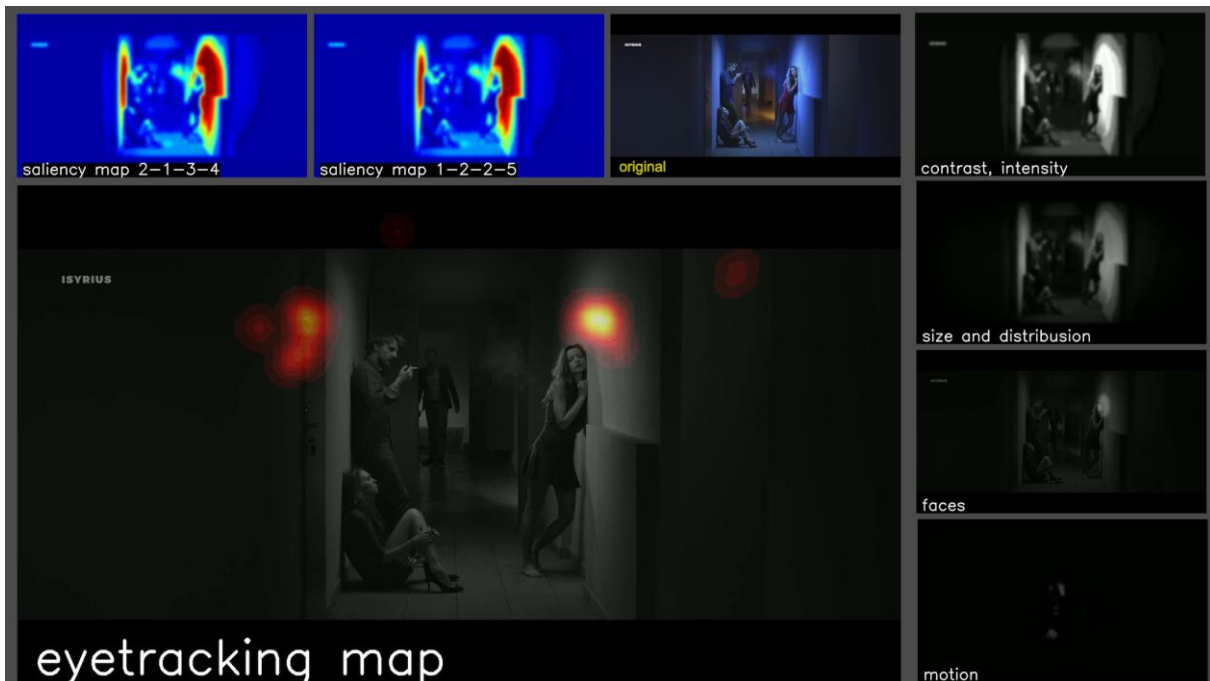
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
0666	Plan bliski, postać na motocyklu, ruch drzew w tle.	Koncentracja wokół twarzy, czasem na iglakach.	Rozpoznana twarz na wprost, iglaki.	Błędnie wyznaczone czarne pasy.	80%



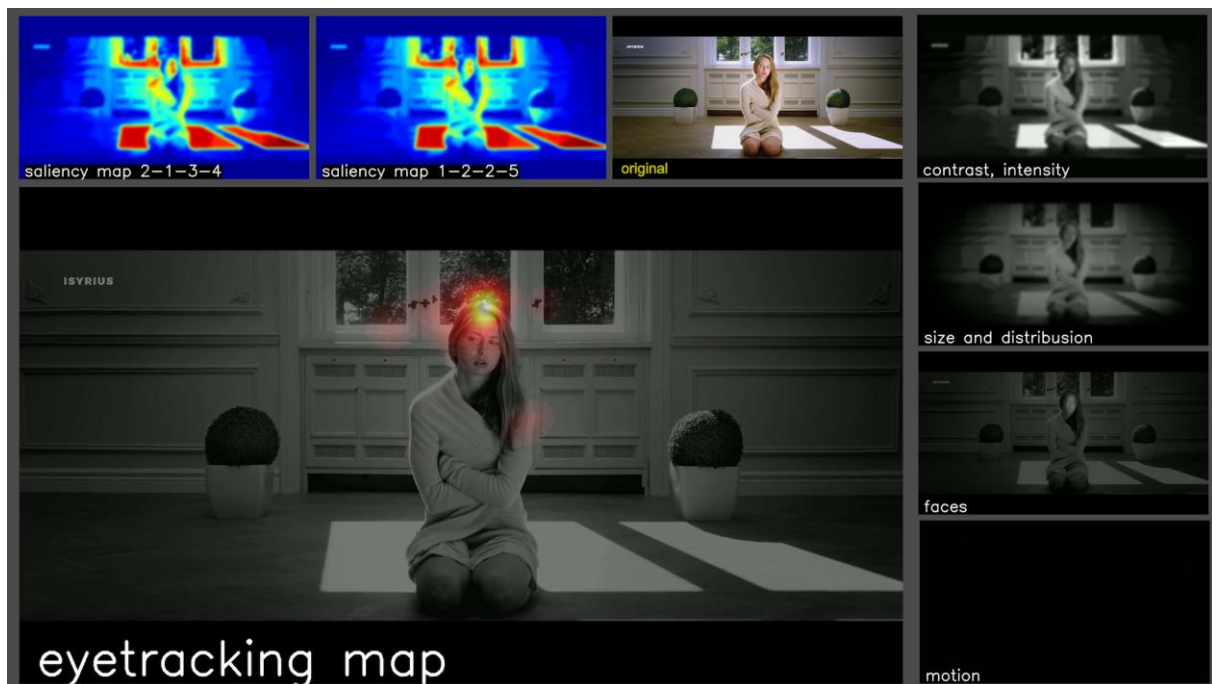
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
0701	Plan ogólny, postacie w całym kadrze, półmrok, łagodny ruch kobiety w centrum.	Koncentracja wokół centrum oraz postaci.	Koncentracja w jasnych obszarach.	Nie wykryto postaci, wykryto tylko niektóre twarze. Pomimo, że postaci są ciemne, powinny one bardziej przyciągać uwagę niż np: kolorowa ściana.	40%



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
0724	Plan ogólny, scena statyczna, postacie w centrum kadru, półmrok, światło na stoliku.	Koncentracja wokół twarzy postaci oraz okna.	Rozpoznana jedna z twarzy. Identyfikacja obiektów, postaci, okna.	Niedostateczna koncentracja na twarzach postaci. Za duża koncentracja na oknie za postaciami.	75%



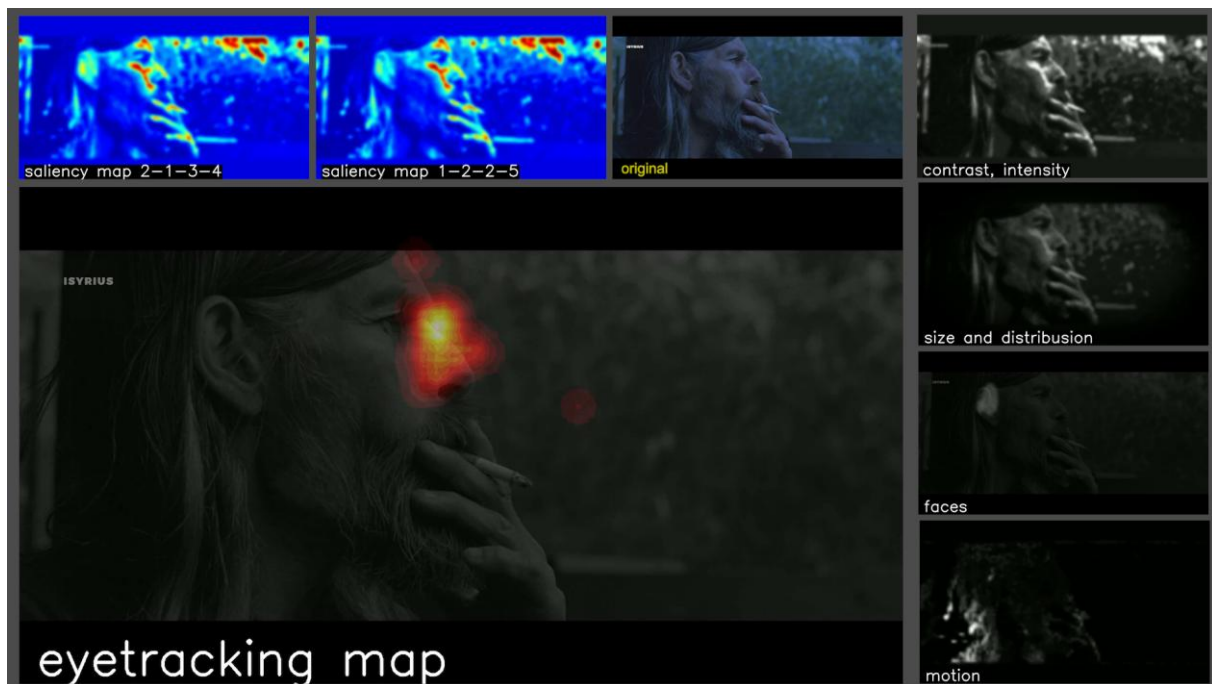
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
0733	Plan ogólny, scena statyczna, postacie w centrum kadru, półmrok, białe światło na ścianach.	Koncentracja w centrum kadru.	Rozpoznana twarz na wprost, identyfikacja obiektów.	Nie rozpoznano wszystkich twarzy, nie rozpoznano postaci, które powinny przyciągać uwagę pomimo tego, że są ciemne.	60%



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
0752	Plan ogólny, scena statyczna, postać kobiety w centrum kadru.	Koncentracja wokół twarzy.	Rozpoznana twarz na wprost (na większości klatek), centrum kadru.	Twarz ma za małą wagę w stosunku do jasnych obszarów na podłodze.	40%



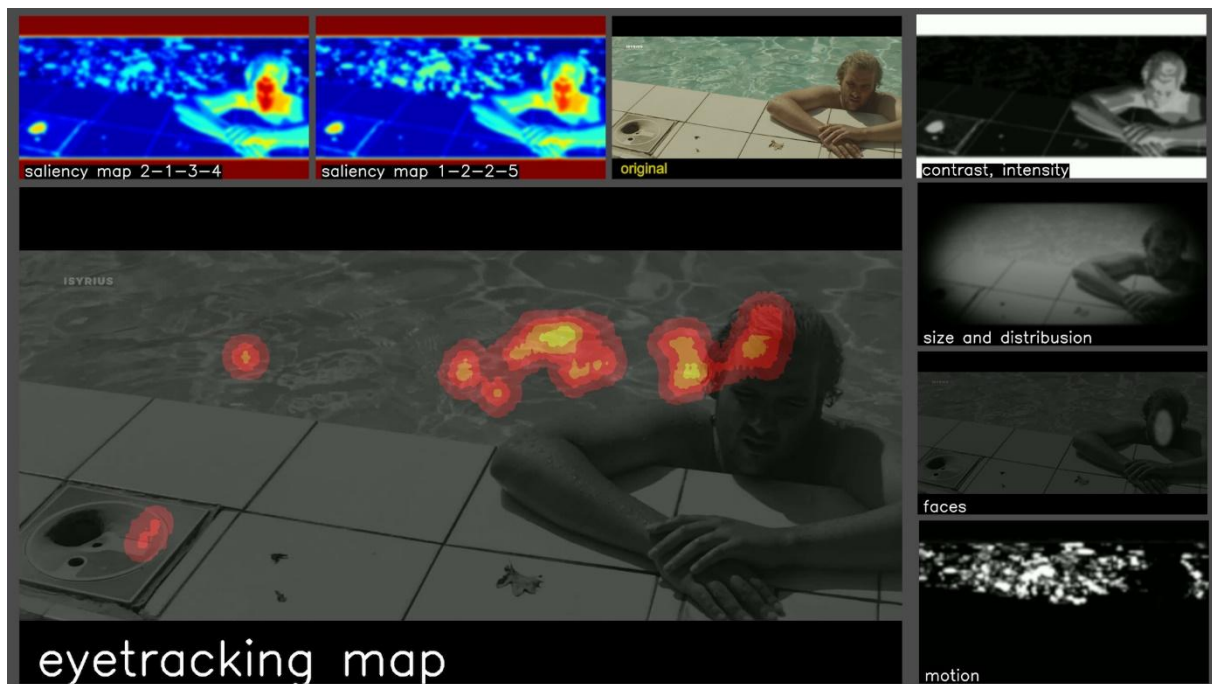
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
0770	Plan ogólny, dwie postaci, ruch ręki z kuflem, dobre oświetlenie.	Koncentracja wokół twarzy (najpierw mężczyzny, później kobiety).	Koncentracja na środku kadru.	Nie rozpoznano twarzy. Kontrast i intensywność zbyt silnie oddziałują na końcową mapę.	50%



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
0787	Zbliżenie, scena statyczna, falowanie włosów, półmrok.	Koncentracja wokół twarzy (na wysokości oczu postaci).	Koncentracja wokół oczu postaci.	Nie rozpoznano twarzy. Błędnie wyznaczone elementy tła.	25%



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
0796	Plan ogólny, scena ruchoma, przeprawa samochodu przez rzekę.	Bardzo krótkie ujęcie. Wzrok koncentrował się w środku kadru.	Identyfikacja obiektów, środek kadru, samochód.	Brak rozpoznania twarzy, kontrast i intensywność wyznaczyły pomarszczoną powierzchnię wody (błędnie). Ruch ma za małą wagę.	40%



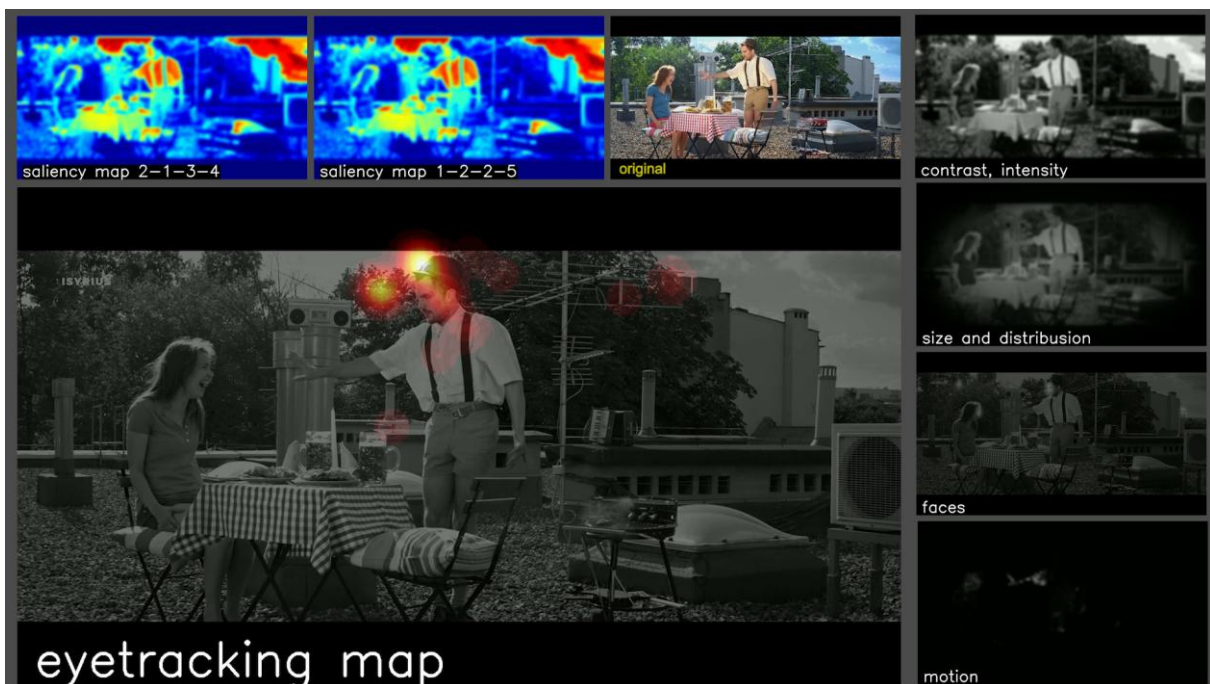
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
0844	Plan średni, scena statyczna, postać z lewej strony, ruch wody w tle.	Wzrok podąża z centrum do twarzy oraz obiektu w lewym dolnym rogu.	Rozpoznana twarz, rozpoznany fragment obiektu po lewej.	Wykryty ruch wody nie jest istotny w tym ujęciu.	90%



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
0873	Plan ogólny, scena statyczna, półmrok, oświetlone postacie w centrum kadru.	Analiza sceny, koncentracja wokół postaci pośrodku.	Koncentracja na środku kadru. Dobrze wykryty kontrast.	Źle wykryta twarz. Twarz powinna mieć większy wpływ na całość.	70%



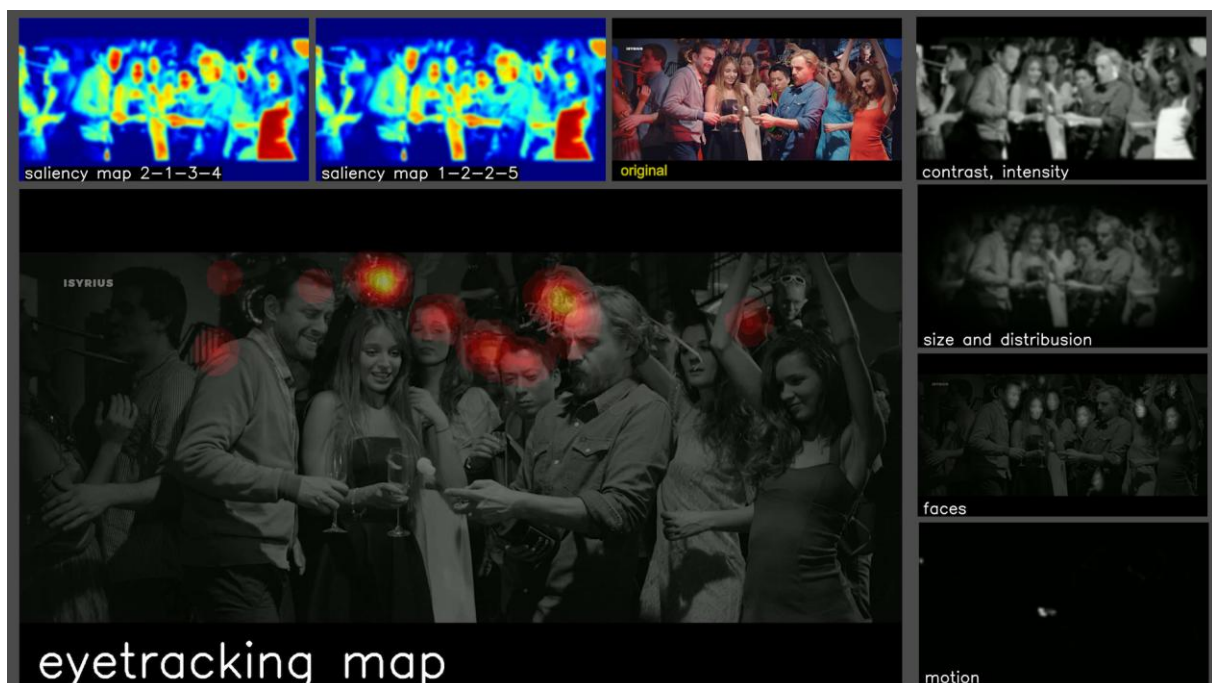
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
0886	Plan ogólny, scena z ruchem, półmrok, światło na postać w centrum kadru.	Koncentracja wokół twarzy postaci na górze.	Koncentracja w górnej części kadru. Wykryty ruch.	Nie rozpoznano wszystkich twarzy. Nie rozpoznano postaci.	25%



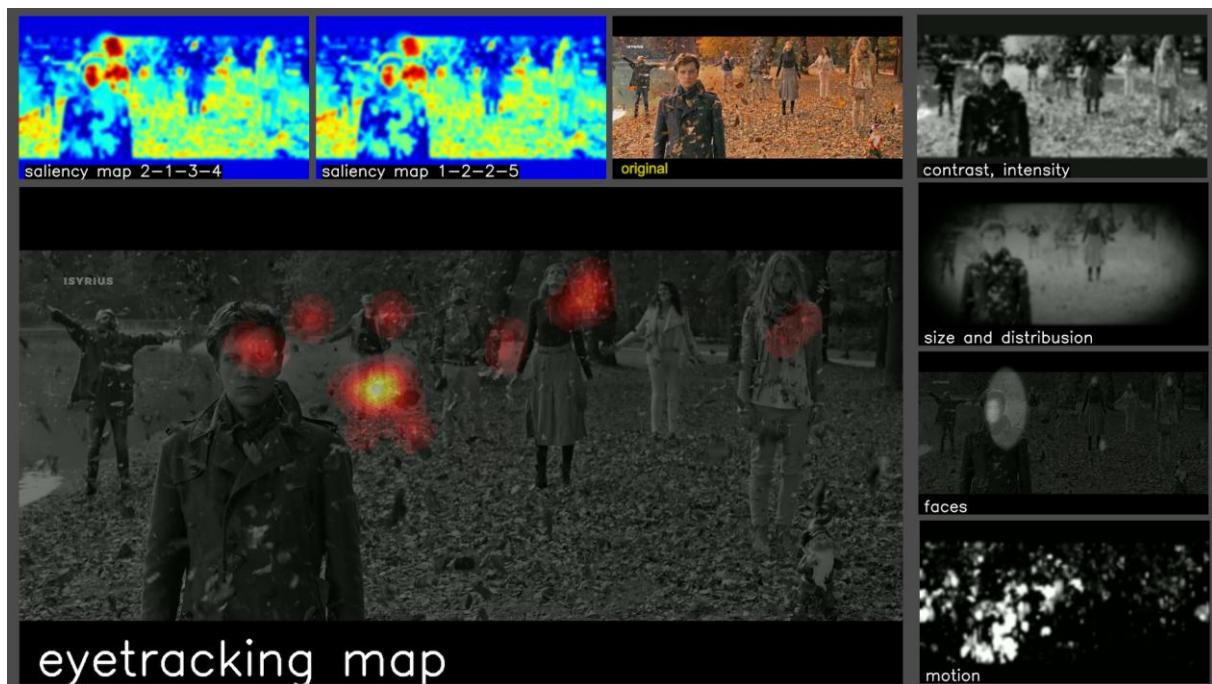
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
0901	Plan ogólny, scena z ruchem, dobre światło, postacie w centrum kadru.	Koncentracja wokół twarzy mężczyzny.	Częściowo rozpoznana twarz. Koncentracja w centrum kadru.	Niebo jest kontrastowe ale w tym przypadku nieistotne. Twarz została przyćmiona przez kontrast.	25%



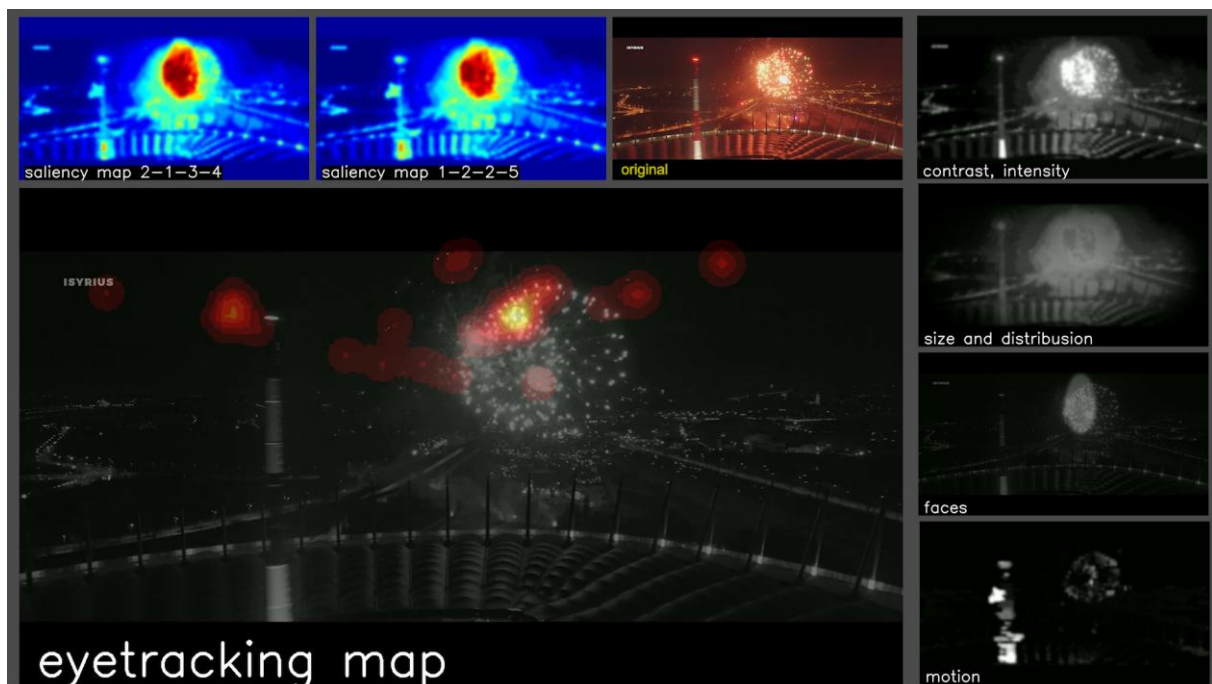
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
0917	Plan amerykański, Scena statyczna, postać w centrum kadru, półmrok, druga postać w tle doświetlona.	Koncentracja wokół twarzy mężczyzny po lewej, mężczyzny w tle oraz obiektu po prawej.	Rozpoznana twarz mężczyzny po lewej, rozpoznany mężczyzna w tle, koncentracja w centrum kadru.	Nie wykryto twarzy mężczyzny na dalszym planie. Nie wykryto obiektu po prawej.	70%



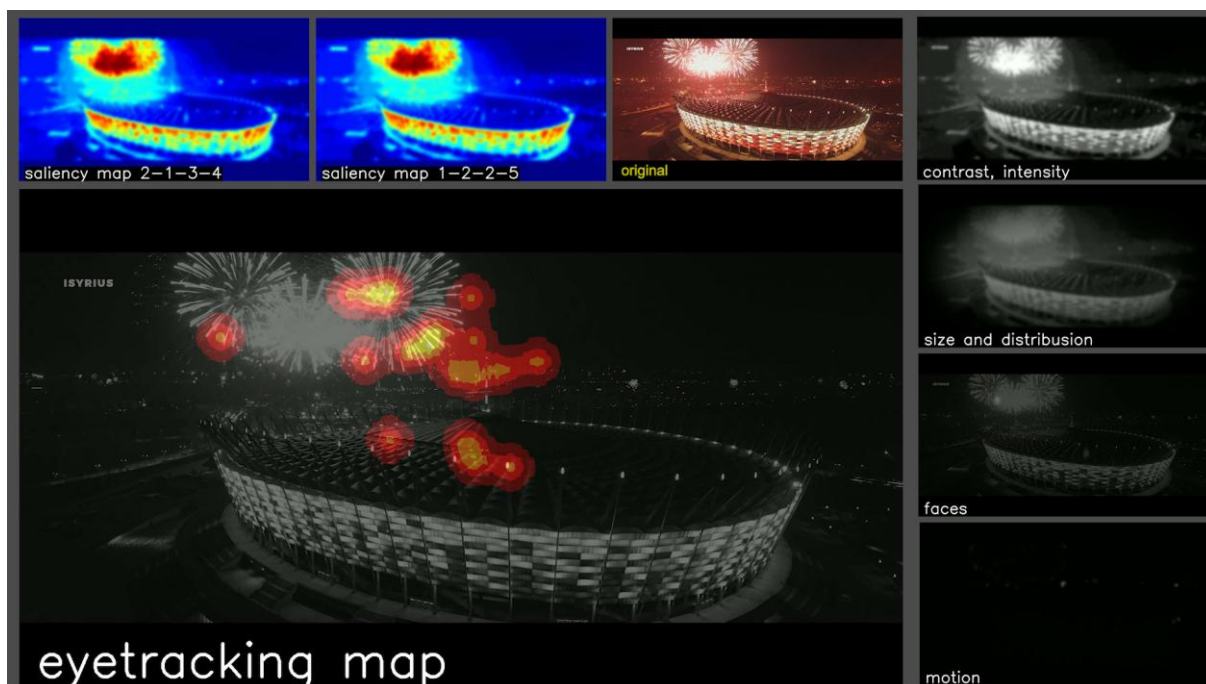
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
0928	Plan amerykański, scena grupowa, postacie w centrum oświetlone, ruch slowmotion korka z butelki.	Koncentracja wokół twarzy centralnych postaci.	Koncentracja w centrum oraz na postaciach.	Błędnie wykryty niebieski element w prawym górnym rogu. Twarze powinny mieć wyższą wagę.	50%



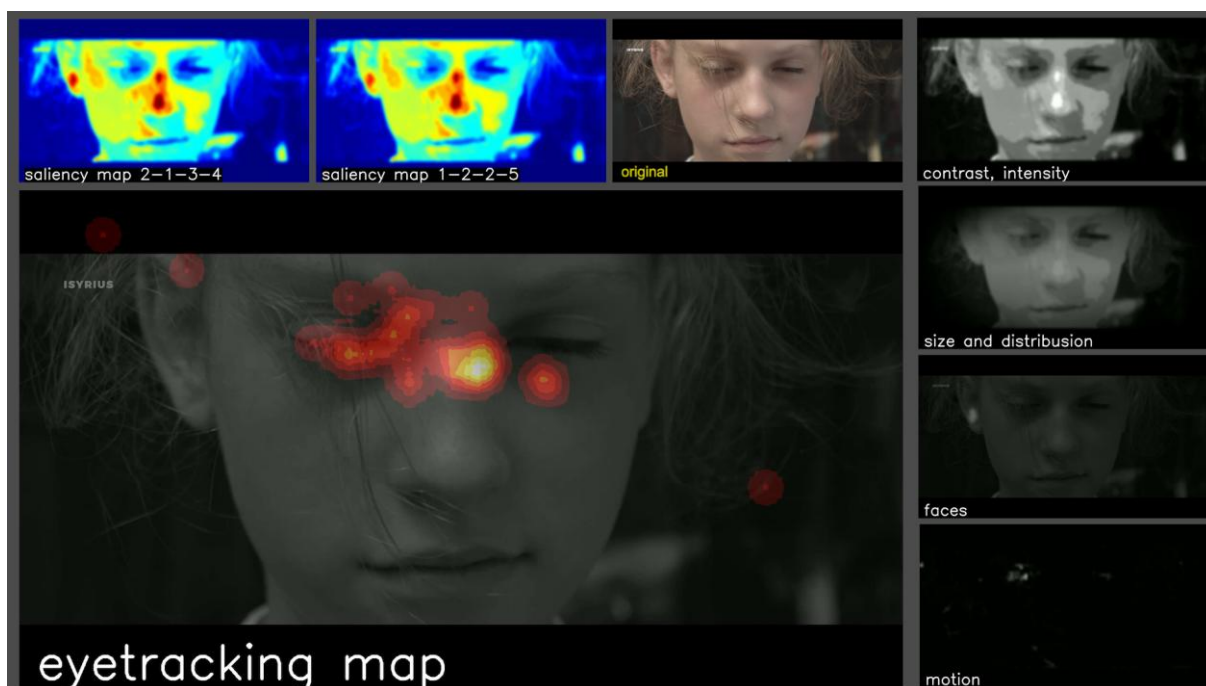
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
0967	Plan ogólny, scena slowmotion, skaczący ludzie, opadające liście.	Koncentracja wokół twarzy postaci oraz wznoszących się liści.	Ruch wznoszących się liści, twarze postaci.	Nie rozpoznano wszystkich twarzy. Za duża koncentracja na ziemi.	50%



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
0991	Plan totalny, wybuchy fajerwerków w nocy.	Koncentracja wokół fajerwerków.	Jaskrawość oraz kontrast dobrze wykryte, koncentracja na środku kadru.	Wykryto błędnie twarz.	90%



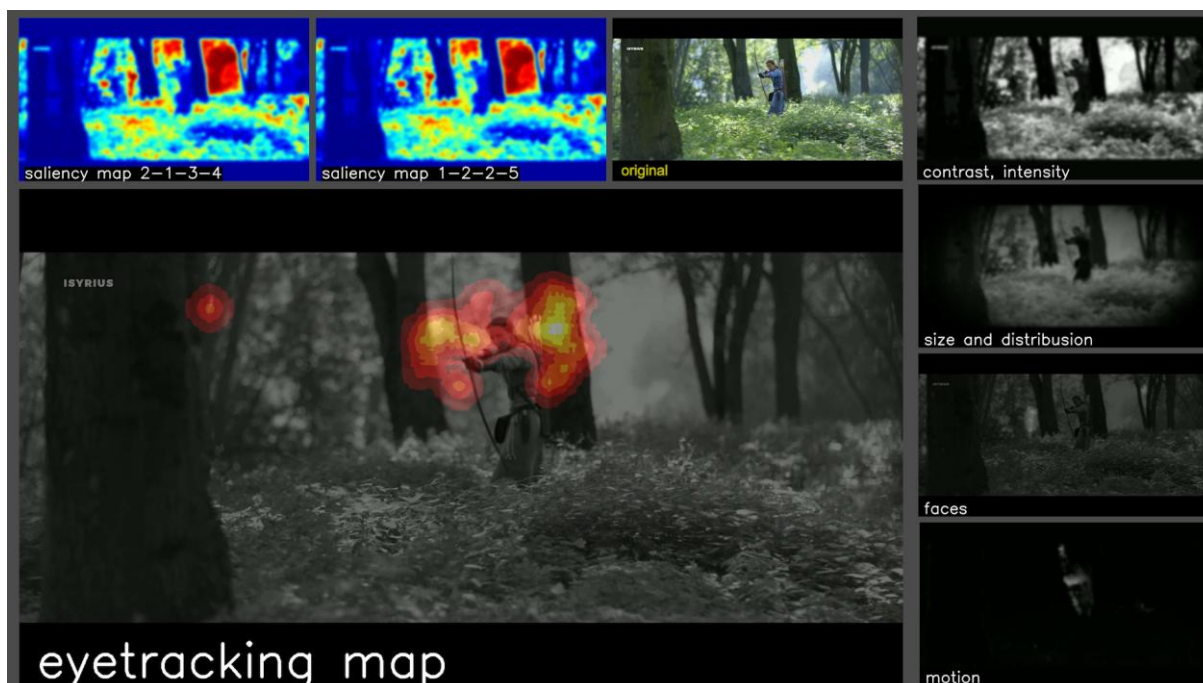
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
1034	Plan totalny, perspektywa, wybuchy fajerwerków w nocy oświetlona bryła stadionu.	Koncentracja w centrum kadru, na fajerwerkach po lewej oraz stadionie.	Fajerwerki mają dużą wagę.	Nie cały stadion przyciąga uwagę.	90%



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
1046	Zbliżenie twarzy, łagodne falowanie włosów.	Koncentracja wokół nosa i oczu postaci.	Dobrze oddzielone tło od twarzy.	Źle wykryta twarz (na uchu).	60%



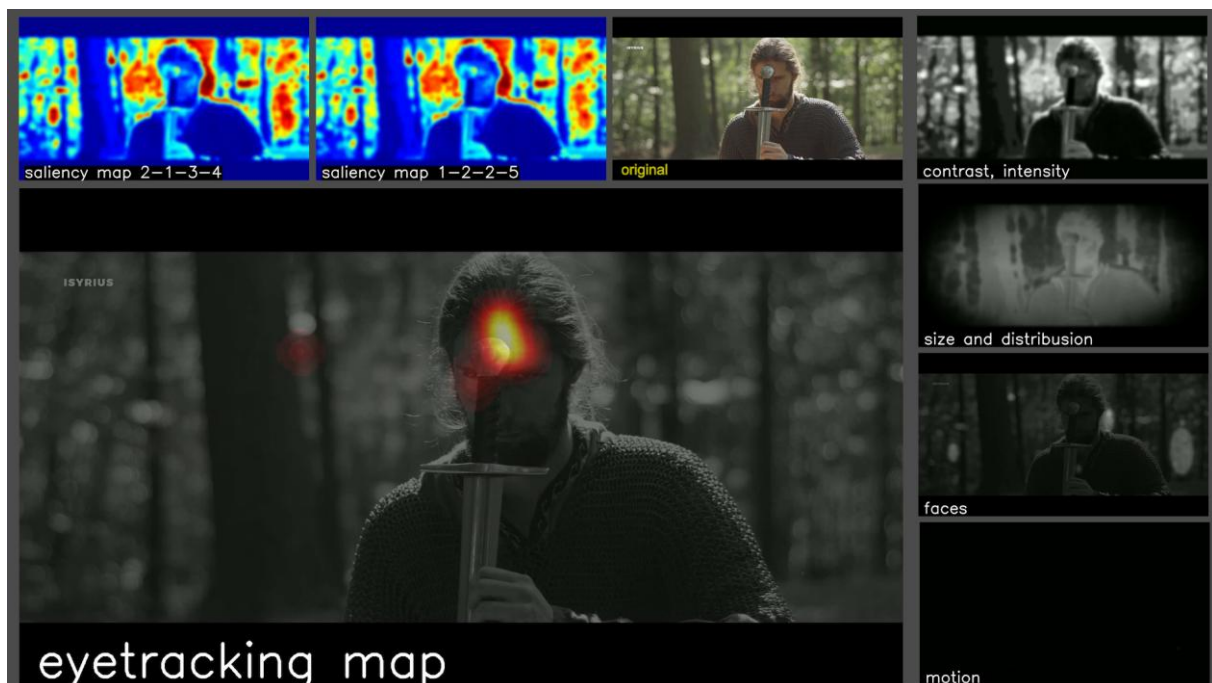
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
1085	Plan pełny, przejazd samochodu w tle.	Koncentracja na twarzy postaci oraz na przejeżdżającym samochodzie.	Centralna część kadru, rozpoznany ruch.	Postać na koniu powinna mieć wyższą wagę niż czerwony samochód.	60%



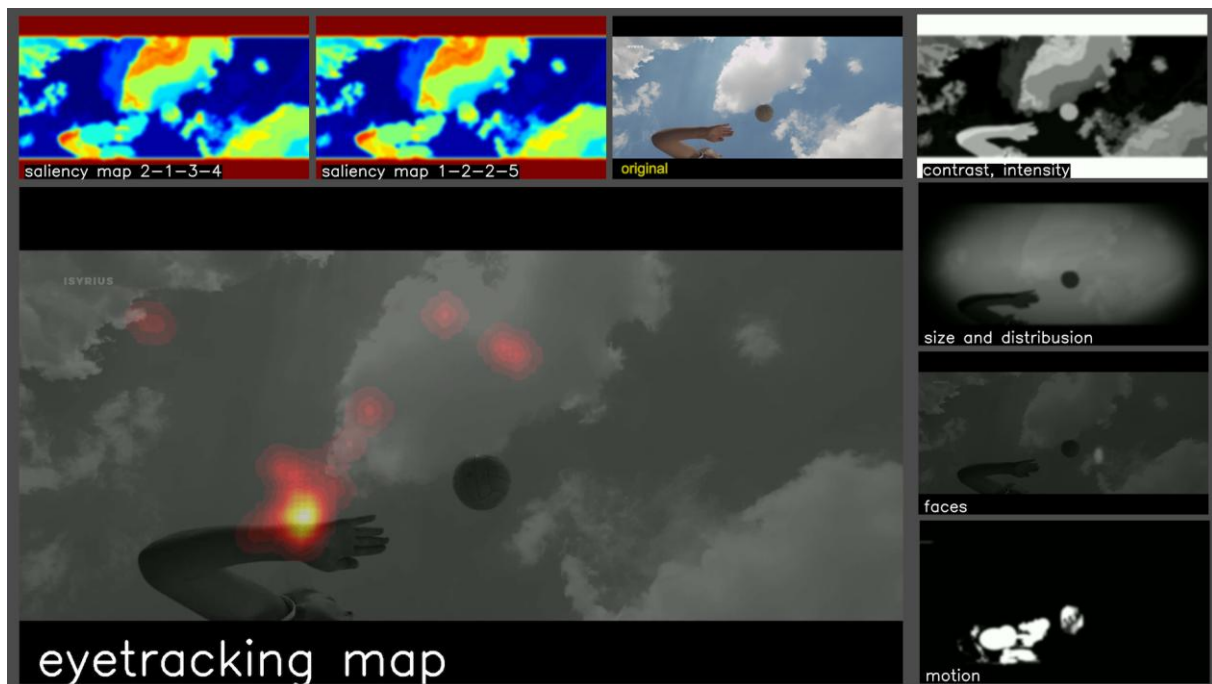
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
1128	Plan pełny, plener, postać strzelająca z łuku w centrum, krótkie ujęcie.	Koncentracja na postaci.	Dobrze wykryty ruch, koncentracja w centrum kadru.	Nie wykryto twarzy, ani postaci. Postać powinna być istotniejsza niż jaste elementy na scenie.	40%



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
1143	Zbliżenie na twarz łuczniczki.	Koncentracja wokół twarzy.	Koncentracja w centrum oraz na elementach ruchomych.	Nie wykryto twarzy. Kontrast oraz intensywność odwracają uwagę od twarzy, która nie jest zbyt kontrastowa w tym ujęciu.	50%



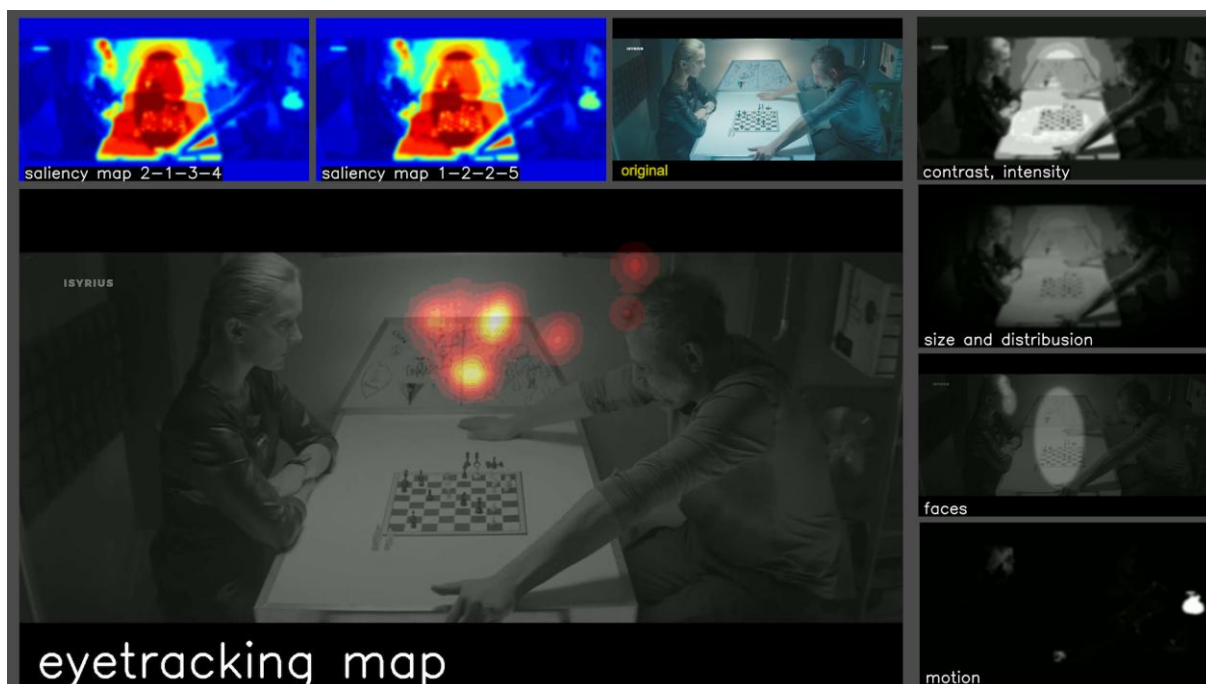
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
1186	Półzbliżenie, scena statyczna, postać w centrum kadru, światło z tyłu.	Koncentracja na mieczu, koncentracja na twarzy postaci.	Koncentracja w centrum kadru.	Nie rozpoznano twarzy (jest zakryta przez miecz). Elementy tła mają zbyt wysoką wagę.	20%



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
1219	Scena dynamiczna, postać wyrzuca piłkę w niebo	Koncentracja najpierw rozproszona, a później na piłce.	Dobrze wykryty ruch.	Kontrastowe elementy tła (białe chmury na niebieskim niebie) przekłamują mapę wynikową. Błędnie wykryta twarz.	50%



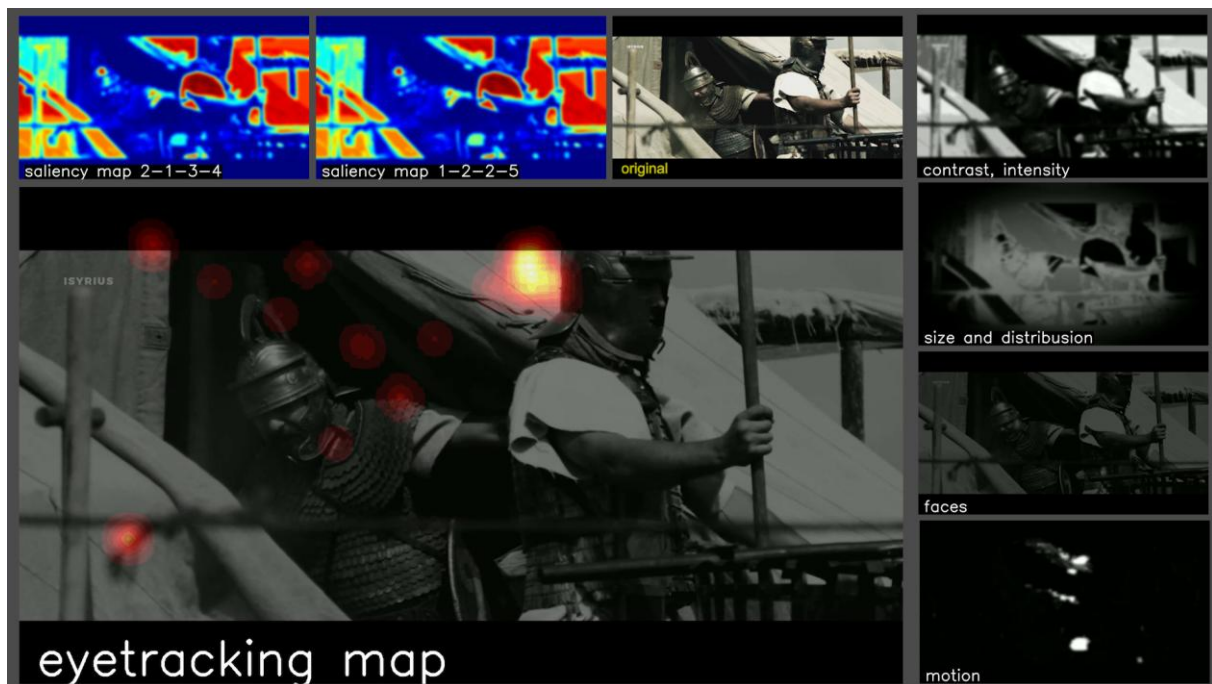
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
1244	Plan średni, scena dynamiczna, przejazd auta przed kamerą	Koncentracja wokół twarzy postaci.	Wykryto ruch. Koncentracja w środku kadru.	Nie wykryto twarzy. Kontrast i intensywność odwracają uwagę od centralnych postaci.	50%



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
1269	Plan średni, scena statyczna, postacie z profile wokół centrum, oświetlony stół z szachownicą.	Koncentracja wokół twarzy postaci oraz na szachownicy.	Koncentracja w centrum.	Błędnie wykryta twarz. Ruch po prawej jest nieistotny.	90%



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
1327	Plan amerykański. Scena z ruchem, postać po prawo wkłada hełm.	Koncentracja wokół twarzy postaci oraz w miejscach ruchomych.	Koncentracja w centrum kadru, ruch. Bardzo dobrze wykryty element po lewej stronie.	Źle wykryta twarz.	90%



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
1389	Plan średni, dwie postacie w centrum, jedna wychodzi z namiotu, duży kontrast świateł.	Skanowanie kadru, poszukiwanie twarzy postaci.	Elementy kontrastowe na gladiatorach.	Nie wykryto twarzy. Kontrast oraz intensywność namiotów i nieba nie powinny przyciągać uwagi.	10%



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
1434	Plan pełny, ruch kamery wokół grupy postaci zajmującej cały kadr.	Koncentracja wokół postaci (nie tylko twarzy).	Koncentracja w środku kadru.	Nie rozpoznano twarzy postaci.	60%



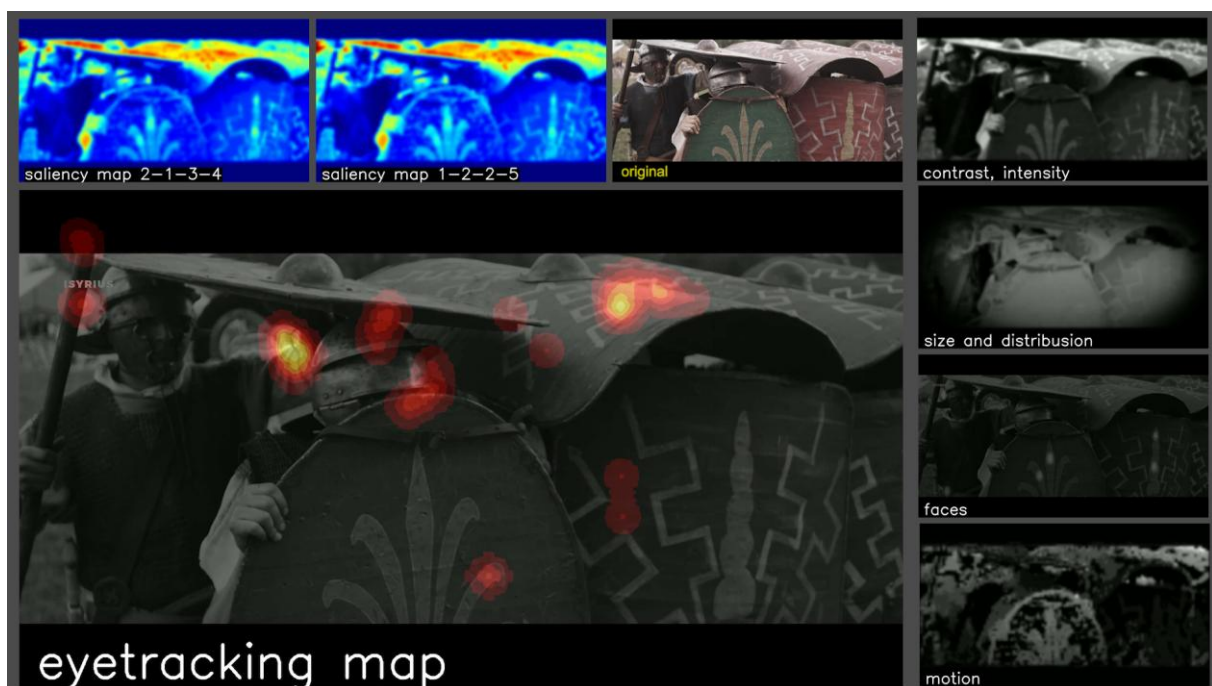
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
1467	Plan pełny, ruch grupy postaci w stronę kamery.	Koncentracja wokół twarzy postaci oraz na ruchomych obiektach.	Rozpoznano ruch oraz koncentrację w środku kadru.	Nie wykryto wszystkich twarzy, a niektóre wykryto błędnie	50%



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
1500	Plan pełny, przemarsz grupy przed kamerą.	Koncentracja wokół twarzy postaci oraz wokół jeźdźców.	Koncentracja w środku kadru i na horyzoncie.	Nie wykryto ruchu koni (który w tym ujęciu był istotniejszy), niebo nie powinno przyciągać uwagi.	40%



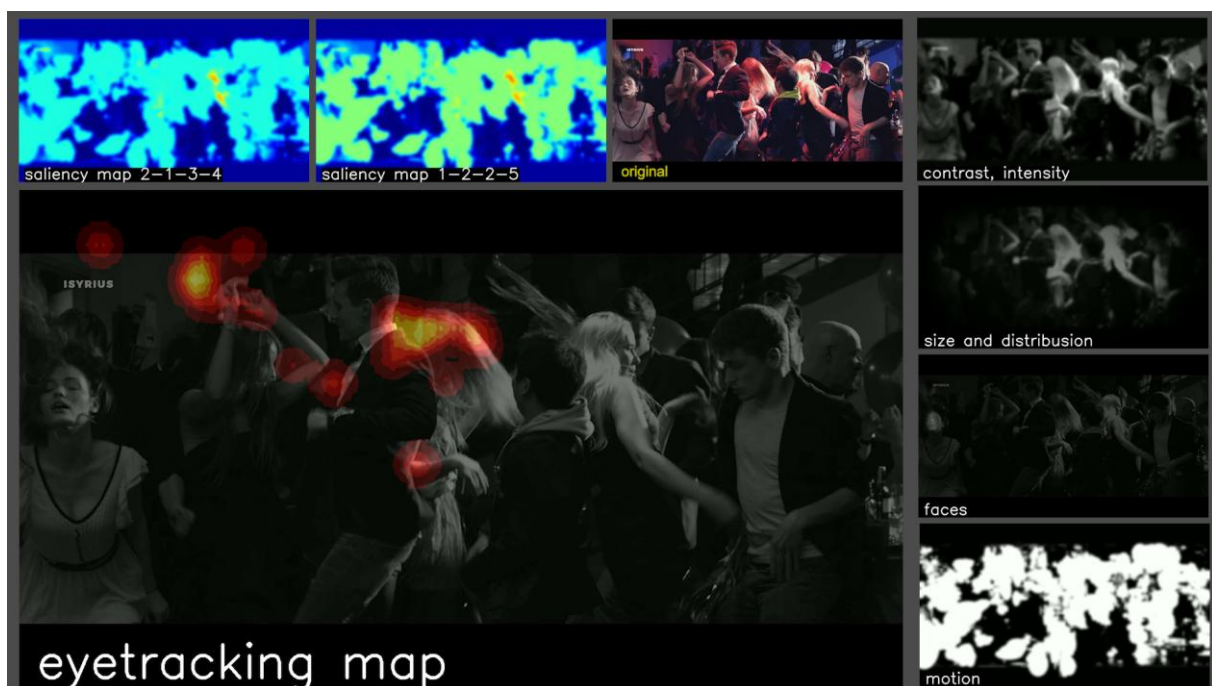
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
1536	Plan bliski, przemarsz grupy przed kamerą.	Koncentracja wokół twarzy postaci.	Wykrycie ruchu, centrum kadru, kontrastowe elementy na gladiatorach.	Nie wykryto twarzy.	50%



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
1546	Plan bliski, bryła z ludzi i tarcz, bardzo krótkie ujęcie.	Koncentracja wokół twarzy.	Wykrycie ruchu, centrum kadru.	Nie wykryto twarzy. Jest to bardzo trudny przypadek, ponieważ twarze są częściowo zasłonięte i zniekształcone przez hełmy (bardzo trudno będzie taki przypadek rozwiązać).	0%



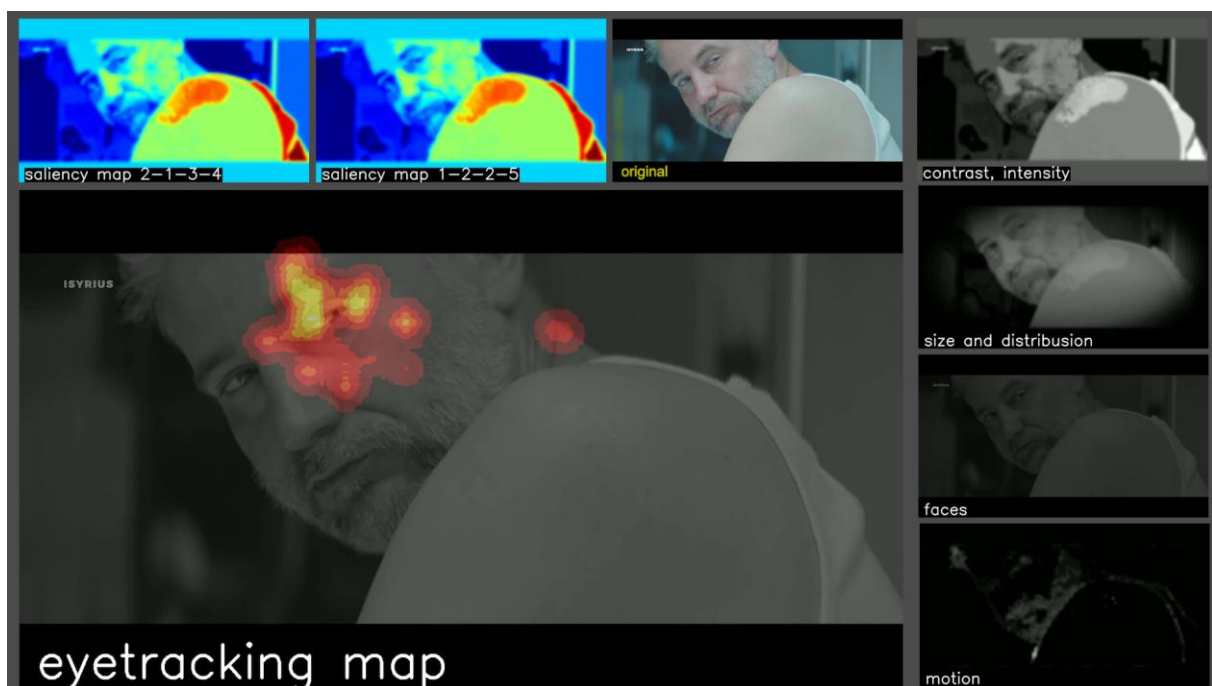
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
1573	Plan pełny, scena dynamiczna, przejazd rydwanu przed kamerą.	Koncentracja wokół twarzy postaci oraz zwierząt.	Wykrycie ruchu, koncentracja w centrum kadru.	Nie rozpoznano wszystkich twarzy. Trawa nie jest istotna.	10%



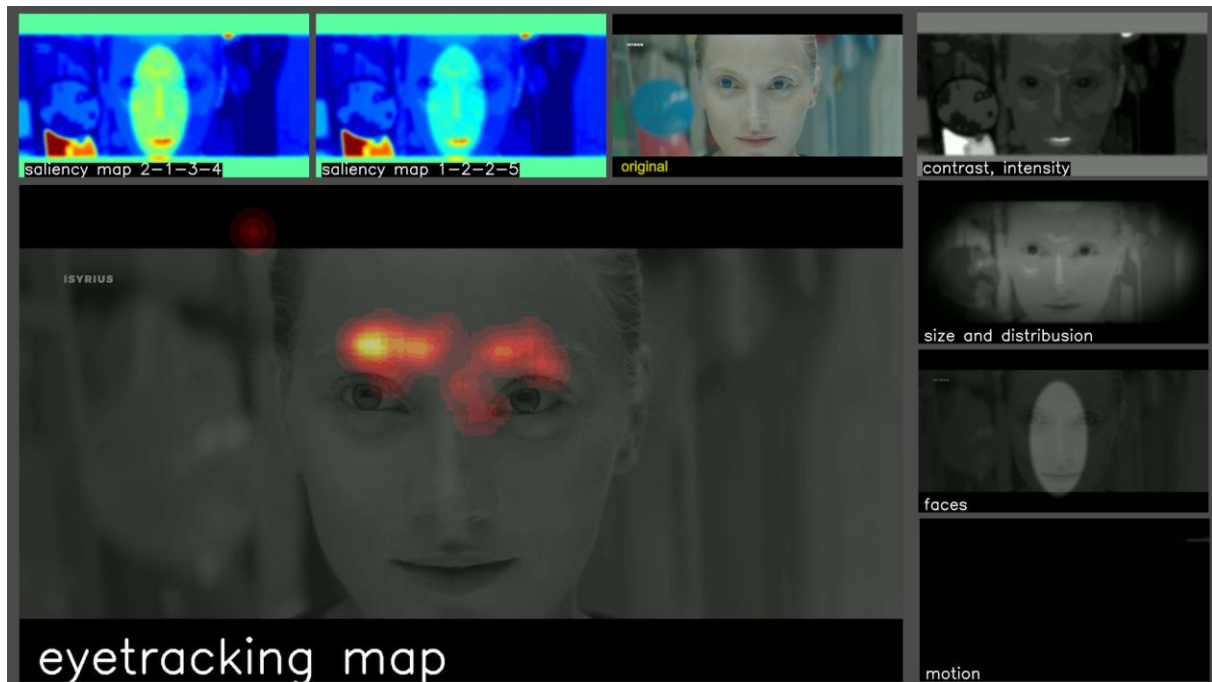
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
1597	Plan amerykański, scena grupowa, taniec, dynamiczny ruch doświetlonych, centralnych postaci.	Koncentracja wokół twarzy postaci.	Rozpoznana twarz na wprost, ruch, kontrast.	Wykryty ruch nie jest aż tak istotny. Ważniejsze są twarze oraz koncentracja w centrum kadru.	50%



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
1624	Plan amerykański, scena grupowa, taniec, dynamiczny ruch włosów, obiekty w centrum.	Koncentracja wokół twarzy postaci, ruch włosów dziewczyny.	Koncentracja w centrum kadru, ruch włosów postaci, kontrast intensywność dobrze zadziałały.	Część twarzy jest błędnie wykryta.	90%



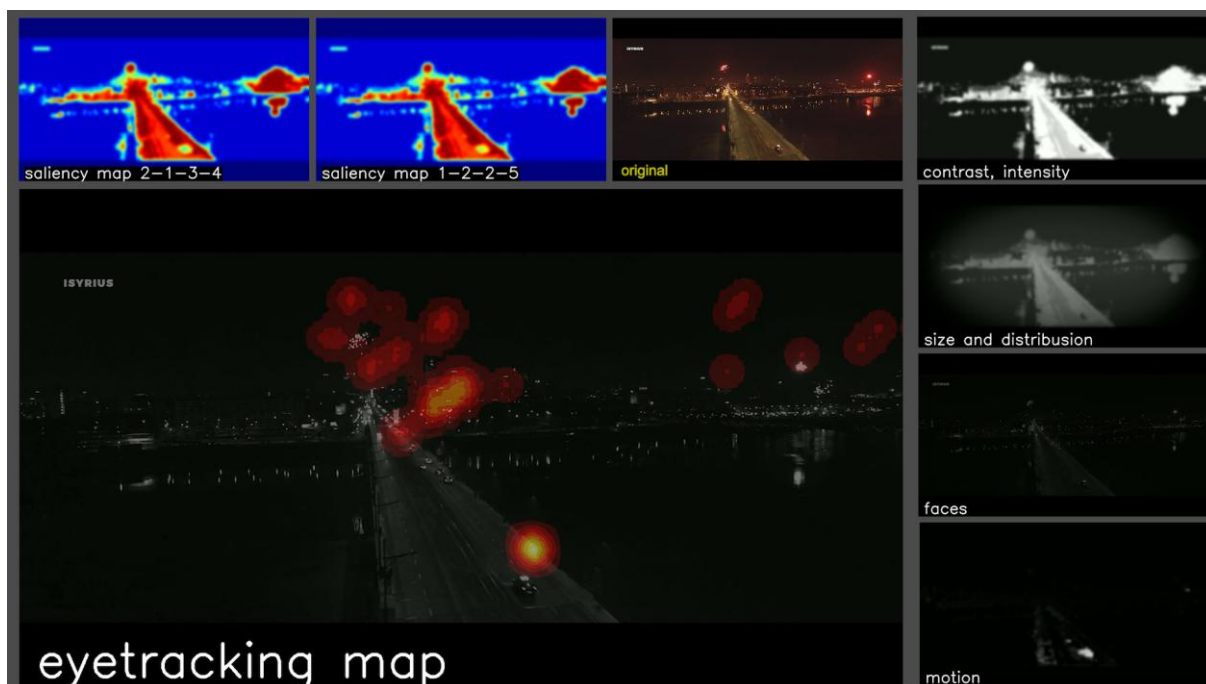
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
1721	Zbliżenie na twarz odchylającej się postaci zajmującej cały kadr.	Koncentracja wokół twarzy.	Oddzielenie postaci od tła.	Nie rozpoznano twarzy. Koszulka za bardzo przyciąga uwagę.	60%



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
1742	Zbliżenie centralne na twarz kobiecą, scena statyczna.	Koncentracja wokół oczu i nosa postaci.	Rozpoznana twarz na wprost.	Kontrastowy element nie jest istotny w przypadku ujęcia tej długości. Kształt twarzy źle wyznaczony.	50%



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
1816	Plan pełny, przejazd kobiety na koniu przed kamerą w poprzek kadru.	Początkowo koncentracja wokół twarzy postaci na koniu, później na twarzy starszej kobiety w oknie.	Koncentracja wokół prawej części kadru.	Błędne wykrycie twarzy, nie wykryto ruchu (ruch slow motion). Ziemia nie powinna przyciągać uwagi.	30%



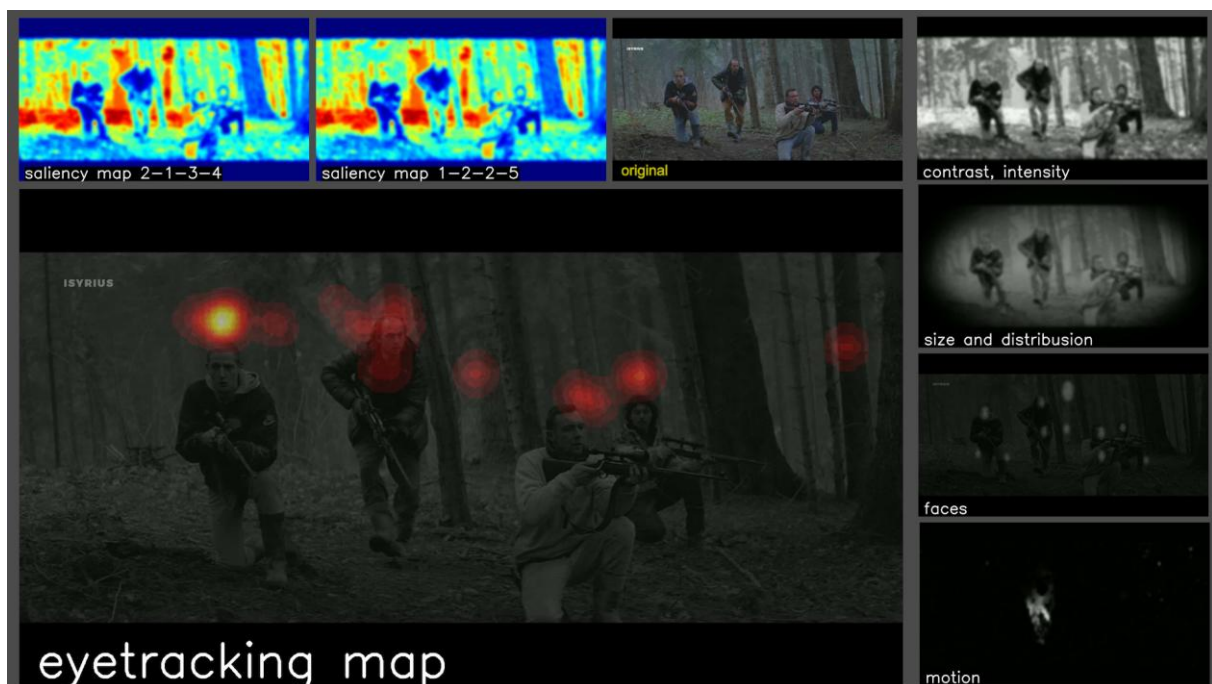
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość	Zgodność
1858	Plan totalny, miasto i most nocą, ruch na moście.	Wzrok skupiony na drodze, na horyzoncie, na światłach, na samochodach.	Dobrze wykryta droga, horyzont, czerwone światło po prawej.	-	90%



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
1908	Plan pełny, plener, jazda kamery za samochodem.	Wzrok koncentruje się w środku kadru oraz na samochodzie.	brak	Wykryty ruch wynika z ruchu kamery, a nie pojazdu. Taki ruch nie przyciąga wzroku. Niebo za mocno przyciąga wzrok.	0%



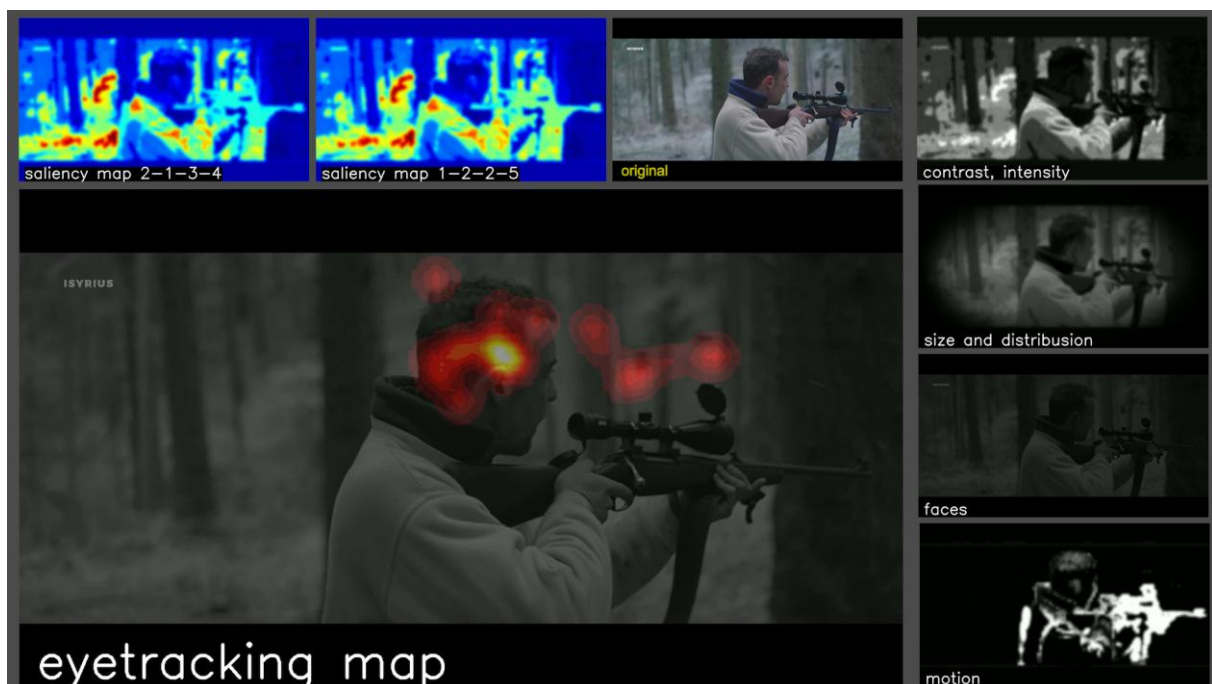
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
1936	Plan bliski. Grupa postaci we wnętrzu samochodu.	Wzrok koncentruje się na twarzach postaci.	Wykryto twarze.	Wykryte twarze mają za mały wpływ na końcową mapę. Wykryty obszar twarzy jest za mały. Śnieg za oknem nie powinien przyciągać wzroku.	10%



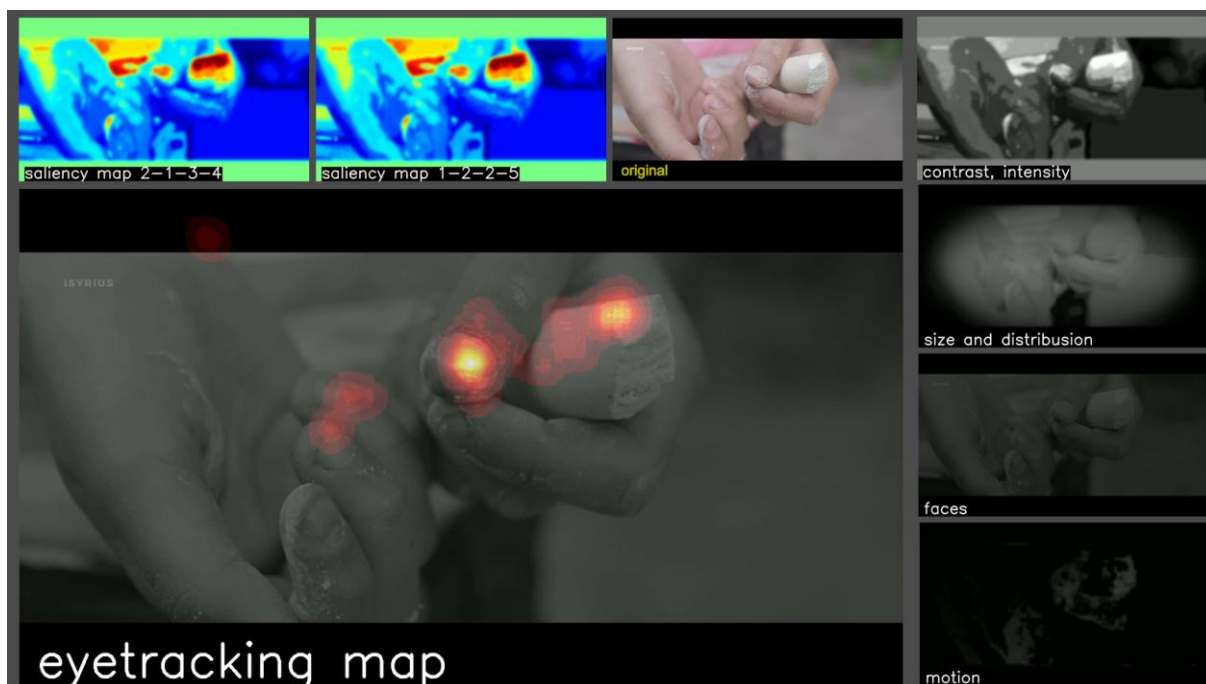
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
1982	Plan amerykański. Grupa postaci z bronią frontem do kamery.	Wzrok koncentruje się na twarzach postaci.	W przypadku mapy 1-2-2-5 twarze zostały wykryte.	Dużo twarzy wykryto błędnie. Kontrast i intensywność nie dają w tym wypadku dobrych wyników.	20%



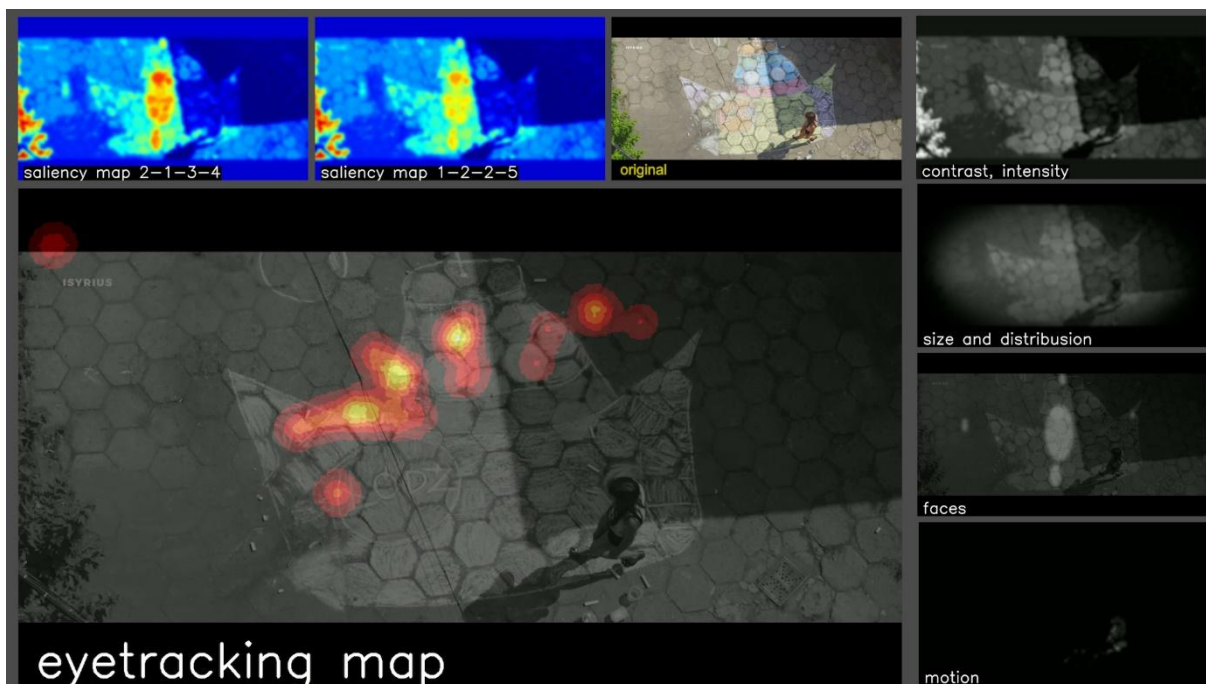
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
2018	Plan średni. Kamera śledzi poruszającą się w poprzek kadru postać z bronią.	Wzrok koncentruje się wzdłuż lufy i na postaci na drugim planie.	Dobrze wykryto ruch przy lufie.	Tło ma zbyt dużą wartość uwagi (w stosunku do lufy).	60%



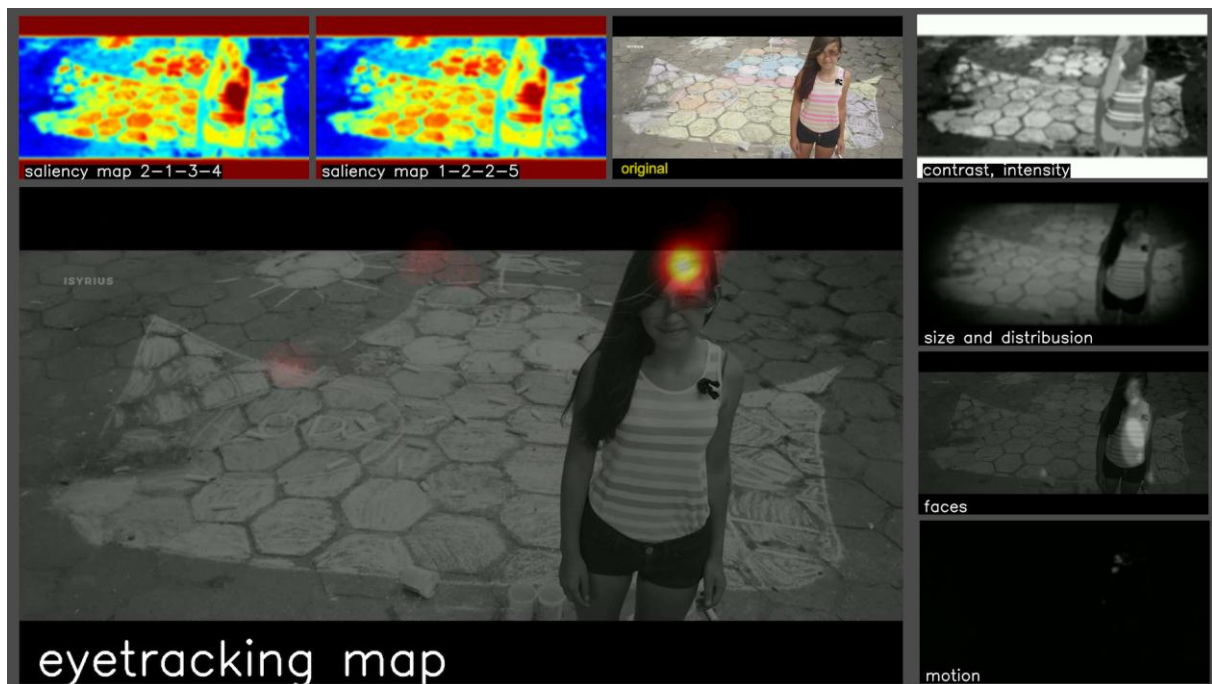
Klatka	Opis sceny	Eyetracking	Trafność	Słabość mapy	Zgodność
2070	Plan średni. Kamera śledzi poruszającą się w poprzek kadru postać z bronią.	Wzrok koncentruje się na twarzy postaci oraz wzdłuż lufy.	Dobrze wykryto ruch.	Wykrywanie twarzy nie działa dla tego przypadku (głowa jest widoczna częściowo od tyłu).	30%



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
2091	Plan detal. Dłonie umazane kredą i kreda.	Wzrok koncentruje się rękach oraz na kredzie.	Częściowo wykryte dłonie i kreda	Wyznaczony jest zbyt duży obszar.	20%



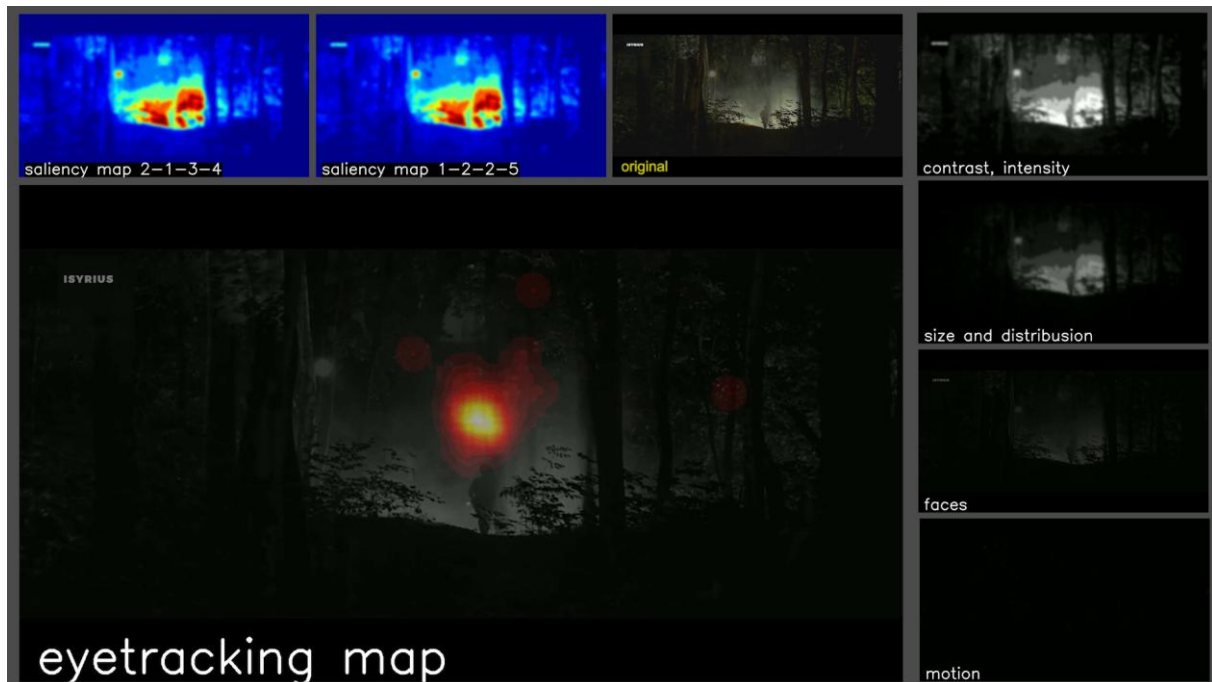
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
2109	Plan daleji z góry. Dziewczynka na tle rysunku na betonowych kostkach.	Wzrok koncentruje się rysunku na ziemi.	Koncentracja w środku kadru.	Błędnie wykryte twarze.	60%



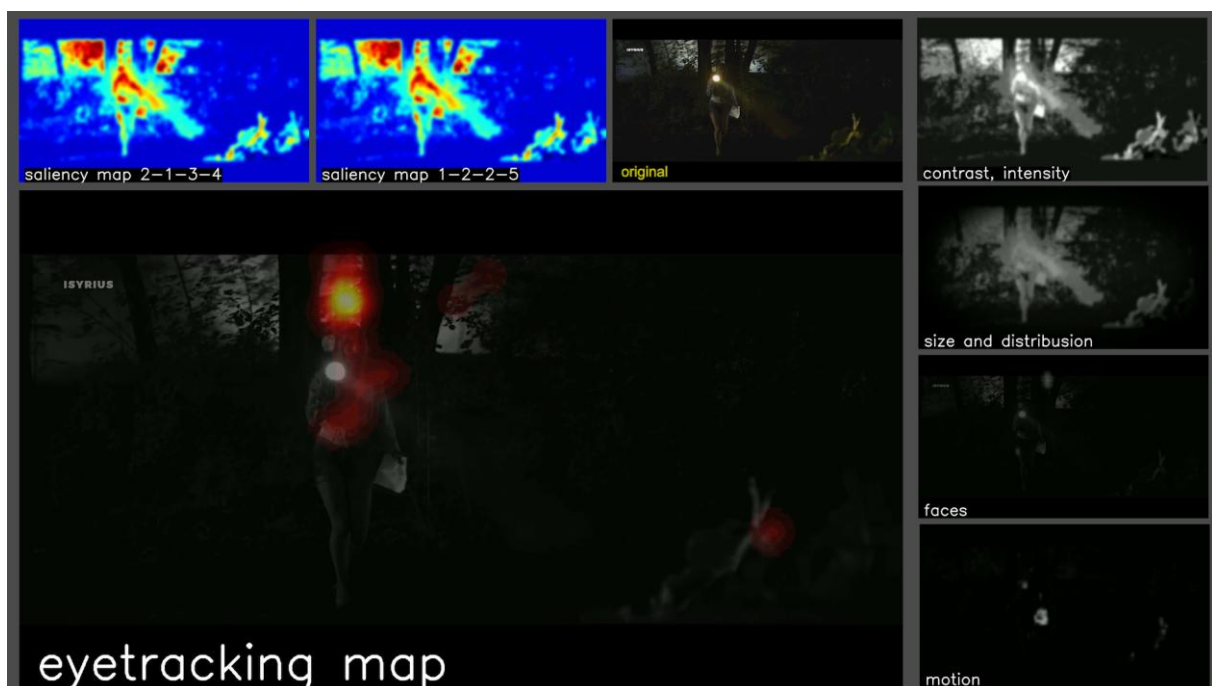
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
2146	Plan półpełny. Dziewczynka na tle rysunku na betonowych kostkach	Wzrok koncentruje się na twarzy i na fragmentach rysunku.	Kolor i intensywność daje dobre wyniki.	Źle rozpoznano twarz, twarz ma za mały wpływ na całą mapę.	60%



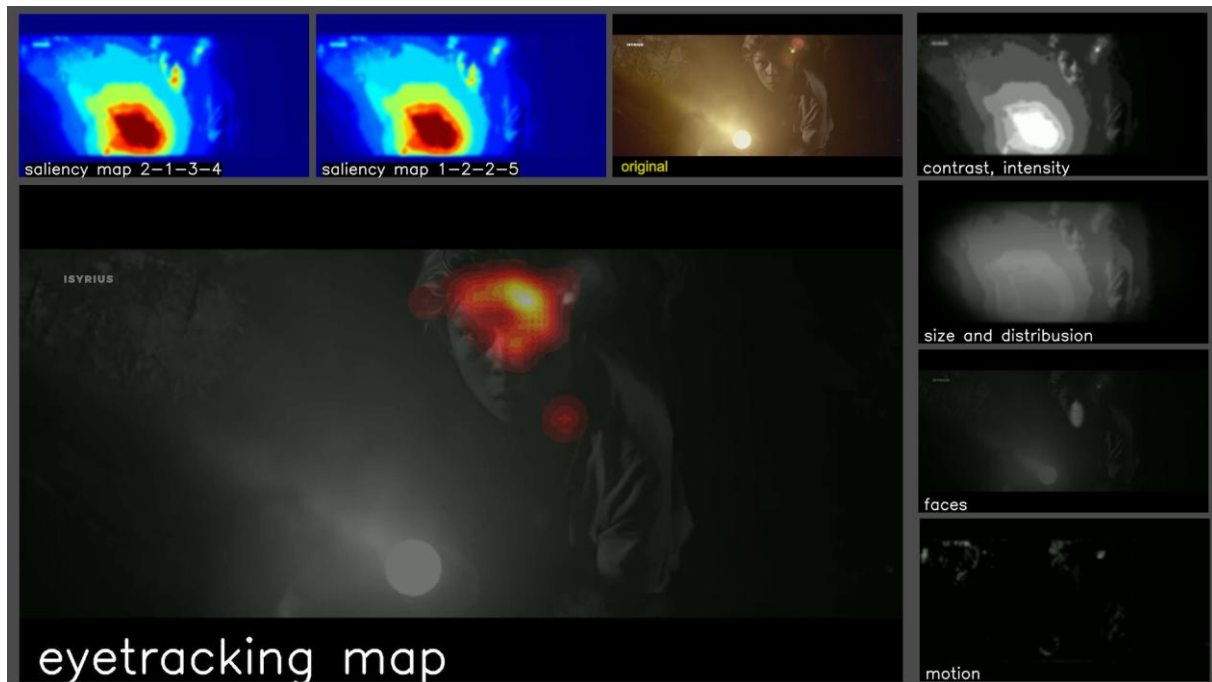
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
2166	Plan półpełny. Scena statyczna. Postacie siedzące w centrum kadru.	Wzrok koncentruje się na twarzach postaci oraz na obiektach na murku.	Wyznaczono sylwetkę mężczyzn oraz obiekty na murku. Wykryto twarze postaci.	Wykryta twarz ma za mały wpływ na końcową mapę. W tym przypadku konewka, aż tak nie powinna przyciągać wzroku.	80%



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
2215	Plan daleki. Postać w centrum kadru. Półmrok.	Wzrok koncentruje się na jasnym obszarze oraz na postaci.	Wszystkie mapy składowe zadziałały prawidłowo.	-	90%



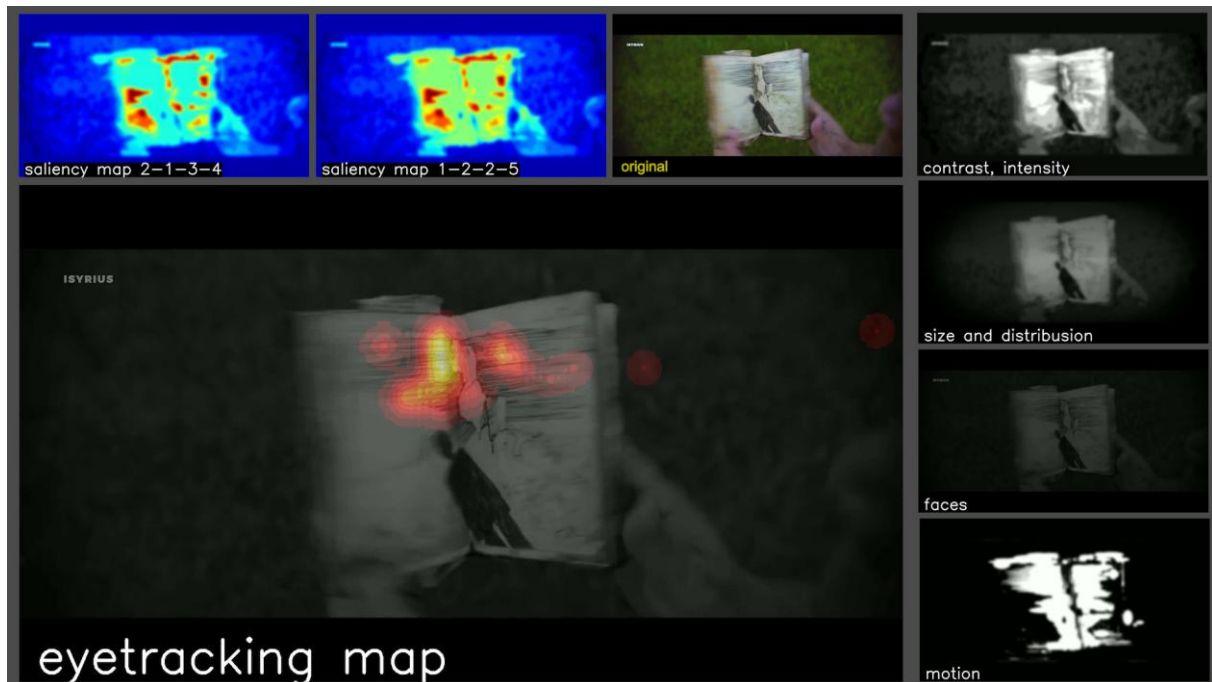
Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
2243	Plan średni. Postać z latarką w centrum kadru. Półmrok.	Wzrok koncentruje się na postaci oraz na twarzy.	Jasność kontrast wyznaczyło odpowiednie obszary. Bardzo dobrze wyznaczony trudny obszar po prawej.	Twarz trudna do rozpoznania.	80%



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
2254	Plan bliski. Postać z latarką w centrum kadru. Półmrok.	Wzrok koncentruje się na twarzy postaci oraz w centrum kadru.	Wyznaczone obszary są poprawne.	Wykryty obszar twarzy jest za mały. Twarz powinna być istotniejsza niż bardzo jasny obszar (światło latarki).	80%



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
2302	Plan daleki. Nałożone efekty specjalne. Ruch kartek książki i postaci.	Wzrok koncentruje się na książce oraz sporadycznie na postaci.	Koncentracja na książce oraz na postaciach (wykryta mapa ruchu)	Jasność i kontrast wyznaczają zbyt duży obszar.	50%



Klatka	Opis sceny	Eyetracking	Trafność mapy	Słabość mapy	Zgodność
2307	Plan bliski. Nałożone efekty specjalne. Dynamiczny ruch przrzućania kartek książki.	Wzrok koncentruje się na książce.	Mapy dobrze oddają obszar istotności (w tej scenie ruch jest istotny).	-	80%

2.7. Wskaźniki korelacji pomiędzy wygenerowaną mapą, a wynikami badania okulograficznego

W celu sprawdzenia skuteczności przeanalizowany został zebrany materiał zawierający dane okulograficzne z wygenerowaną mapą uwagi pod kątem współczynnika korelacji (LCC – Linear Correlation Coefficient).

Współczynnik korelacji liniowej określa zależność pomiędzy dwiema liczbami losowymi.

W naszym przypadku działamy na dwuwymiarowej tablicy liczb losowych o wymiarach wczytanego pliku wideo. Pary liczb składają się z odpowiadających sobie wartości pikseli (o zakresie od 0 do 255) mapy uwagi i mapy z danymi okulograficznymi. Korelacja daje nam informację o tym, jak zmiany w mapie z danymi z okulografu wpływają na mapę uwagi.

Korelacja może być słaba, bądź silna, dodatnia lub ujemna (czyli przyjmować wartości z zakresu $<-1,0;1,0>$). Korelacja dodatnia oznacza, że zmiany w mapie z danymi okulograficznymi prowokują takie same zmiany w mapie uwagi (co oznacza, że np.: czarne obszary w jednej mapie występują w tych samych miejscach, co w drugiej mapie). Korelacja ujemna oznacza, że zmiany w jednej mapie prowokują przeciwne zmiany w drugiej mapie (czarne obszary w jednej mapie występują w tych samych miejscach, co białe obszary w drugiej mapie).

Poniżej znajduje się obraz, na którym po lewej stronie znajduje się mapa uwagi, którą porównujemy z mapą po prawej stronie (mapą z danymi okulograficznymi).

LCC obliczane jest dla każdej klatki filmu. Na obecnym etapie prac, gdzie mapa uwagi ma niestabilną formę pomiędzy kolejnymi klatkami, LCC nie daje jeszcze wiarygodnych wyników.



Rys. Zestawienie monochromatycznych map uwagi do wyliczenia korelacji LCC dla pojedynczej klatki filmu. Wynikowa Saliency map z uwzględnieniem wszystkich składowych, schemat wag: 3-2-2-1.

3. Podsumowanie i konkluzje

Badanie dostarczyło istotnych spostrzeżeń w obu grupach postawionych celów badawczych.

3.1. Przetestowanie skuteczności algorytmów budujących mapę uwagi – konkluzje

Składowa mapy *contrast, intensity* generowała dobre wyniki w przypadku bardzo ciemnych ujęć z elementami jasnymi np. klatki nr.: 1046, 1085, 1624, 1858, 2215, 2243, 2254. W przypadkach bardziej jednolitych obrazów (mała różnica intensywności pomiędzy obiektami) nie wpływała istotnie na wyznaczenie mapy uwagi. Często jasne elementy tła błędnie przyciągały uwagę.

Składowa mapy *size and distribution* praktycznie w przypadku każdego z ujęć prawidłowo zawężała obszar przyciągający uwagę. Wydaje się jednak, że ta składowa wyznacza obszary przyciągające uwagę dość ogólnikowo i schematycznie. Mapa powinna mieć większy wpływ na końcową mapę uwagi.

Składowa mapy *faces* wyznaczała obszary, w których znajdowała się ludzka twarz. W momencie, gdy metoda prawidłowo wykryła twarz – wyznaczony obszar praktycznie w każdym przypadku zgadzał się z danymi z okulografu. Metoda ta, często nie była w stanie wykryć twarzy (np.: gladiatorzy w zbrojach) lub wykrywała je w miejscach, w których ich w rzeczywistości nie było. Również wyznaczony owal twarzy powinien zajmować większy obszar.

Składowa mapy *motion* wyznaczała obszary, w których wystąpił ruch. Składowa ta działała bardzo dokładnie. Jedyną wadą jest to, że składowa uwzględniała ruch, który powstał na skutek ruchu kamery. Ruch ten nie przyciągał uwagi widza, więc wyznaczone obszary nie zawsze są prawidłowe.

3.2. Analiza przypisywania wag algorytmom w tworzeniu mapy uwagi – konkluzje

Koncepcja przypisywania wag poszczególnym metodom składowym budującym wynikową mapę uwagi miała przetestować zasadność parametryzacji tych wag w wyliczaniu mapy uwagi w zależności od dominującej formuły wizualnej w badanym materiale filmowym. W konsekwencji ma uzasadnić specyficzną funkcjonalność aplikacji CinemaVision SaliencyMap polegającą na możliwości wymuszenia przydziału wag poszczególnym składowym metody przez użytkownika w celu jak największego prawdopodobieństwa zgodności wygenerowanej mapy z preferencjami potencjalnego widza.

Badanie pozwoliło zaobserwować kilka istotnych reguł uzasadniających zastosowanie możliwości parametryzacji istotności składowych algorytmów metody względem konkretnej formuły wizualnej kadru:

- Gdy w ujęciu jest obecna twarz to jest ona elementem najsilniej przyciągającym uwagę.
- W przypadku ujęć z szybko poruszającymi się obiektami to one są najważniejsze (chyba, że znajduje się w ujęciu twarz – wtedy ruch jest na drugim miejscu).
- Składowa mapy *contrast, intensity* powinna być bardziej ograniczona przez mapę *size and distribution*.

- Zaobserwowano, że zwykle wzrok koncentruje się w środku kadru (w szczególności, gdy nie ma w ujęciu twarzy oraz szybko poruszających się obiektów), więc podwyższona waga składowej *size and distribution* w większości przypadków poprawia skuteczność wygenerowanej mapy uwagi.
- Zwykle w ujęciach dynamicznych najważniejszy jest ruch. W ujęciach statycznych twarz. W przypadku braku twarzy – kontrast. W przypadku dużego kontrastowego obiektu – wzrok kieruje się ku środkowi ekranu. Są od tego wyjątki.
- Nie wszystkie poruszające się obiekty przyciągają wzrok. Zależy to od ich wielkości, ruchu kamery oraz umiejscowienia w kadrze.
- W scenach zbiorowych ruch gra małą rolę w przyciąganiu uwagi (uwaga jest rozproszona lub kieruje się do centrum).
- Gdy widzowie przyglądają się twarzy, najpierw ich wzrok skupia się w jej centrum, natomiast gdy ujęcie jest dłuższe – zaczynają się przyglądać poszczególnym częściom twarzy: oczom, ustom, nosom.
- Szybkość ruchu ma znaczenie. Wolny ruch ma dużo mniejszy wpływ na uwagę.
- Przy wyraźnym podziale ujęcia na pierwszy i drugi plan, ruch na dalszym planie nie jest tak istotny jak statyczne postaci na pierwszym planie.
- Gdy porusza się cała postać uwaga skupia się w okolicach jej głowy, a nie obejmuje całej postaci.

3.3. Porównanie mapy uwagi z wynikami badania okulograficznego – konkluzje

- Długość ujęcia istotnie modyfikuje mapę uwagi obserwatora – im dłuższe ujęcie, tym bardziej rozstrzelona jest uwaga.
Porównując mapę uwagi z danymi z okulograficznymi można zauważyć, że długość ujęcia także wpływa na to jak jest ono oglądane. Najpierw widz patrzy w konkretne miejsce (akcja, twarz, kontrast, środek), natomiast później zaczyna on badać scenę, oglądać inne rzeczy. Mapa uwagi nie bierze pod uwagę długości ujęcia – zawsze wyznacza „to pierwsze”, najbardziej prawdopodobne miejsce, w które spojrzy się widz. Sytuacja nie ma miejsca w przypadku sceny z dynamiczną akcją – wówczas to ruch jest kluczowym czynnikiem determinującym wodzenie wzrokiem.
- Wyniki okulograficzne są opóźnione w czasie w stosunku do saliency map o 5 do 7 klatek w pierwszej sekundzie po zmianie ujęcia. Materiał badawczy zbudowany był z wielu ujęć często bardzo krótkich i zmieniających się dynamicznie. Montaż był na ostro i często sąsiadujące ujęcia różniły się dramatycznie kolorystyką, planem, dynamiką. Z dużym prawdopodobieństwem wzrok badanych z pewną bezwładnością reagował na częste zmiany – stąd opóźnienie.

4. Ocena przydatności metody komputerowego generowania mapy uwagi dla materiału filmowego do badania preferencji widzów

Komputerowe generowanie mapy uwagi dla materiału filmowego umożliwia oszacowanie rozłożenia w kadrze miejsc wpływających na uwagę widza. Wyznaczenie wspomnianych miejsc pozwala na usprawnienie oraz przyspieszenie etapu produkcji filmowej. Wyznaczone obszary przyciągające uwagę pokazują, czy zaprojektowana scena kieruje uwagę widza w taki sposób, w jaki zostało to zaplanowane.

Metoda komputerowego generowania mapy uwagi bazuje na mechanizmie uwagi człowieka, który uwzględnia nakładanie się na siebie różnych rzeczy przyciągających uwagę w różnym stopniu (z różnymi wagami). Jak zaobserwowano w badaniu, różne rzeczy przyciągają uwagę w zależności od charakterystyki sceny, natomiast analizując dane okulograficzne zaobserwowano, że widzowie skupiają wzrok w zbliżonym obszarze (dane z badania okulograficznego nie różnią się zasadniczo dla indywidualnych osób). Pozwala to stwierdzić, że możliwe jest stworzenie uniwersalnego modelu generującego mapę uwagi.

Opracowana metoda pozwala wykryć istotne z punktu widzenia zamierzeń twórców elementy obrazu np.:

- w filmach reklamowych oszacować czy kluczowy przekaz reklamowy jest dostatecznie zauważalny,
- w produkcji filmowej oszacować, czy błędy w scenie zostaną zauważone przez widza.

Metoda pozwala również uwidocznić miejsca, które potencjalnie mogą zwrócić uwagę widza, a których nie dostrzegliśmy (np.: kontrastowy element w kadrze) i nie chcemy aby zwracały one na siebie uwagę.

Generalnie pomimo zauważalnych niedoskonałości metody, głównie w metodzie identyfikacji twarzy należy uznać ją za rozwojową i dobrze rokującą. Główne akcenty doskonalące powinny być położone na:

Wzbogacenie bazy identyfikacyjnej twarzy pod względem rozpoznania fizjologicznego,

- wzbogacenie bazy identyfikacyjnej twarzy pod względem kątów względem kamery,
- udoskonalenie metody analizy wielkości i rozłożenia obiektów w kadrze w celu eliminacji schematyzmu koncentrowania się głównie w środku kadru i niskiego progu detekcji obiektów,
- zbudowania metodyki przypisywania wag do algorytmów w budowaniu wynikowej mapy uwagi w zależności od dominującego typu wizualnego badanego materiału,
- zoptymalizowanie programistyczne metody w celu eliminacji generowania błędnych klatek mapy uwagi w złożonych materiałach cechujących się dynamicznym montażem i krótkimi ujęciami.