



ROMAN ROSIEK<sup>1</sup>, JOANNA GANCZAREK<sup>2</sup>, KAROLINA PIETRAS<sup>3</sup>

## Analiza stopnia złożoności obrazu w eyetrackingowych badaniach dzieł sztuki

### Analysis of Image Complexity in Eyetracking Studies of Visual Art

<sup>1</sup> Doktor habilitowany profesor UP, Uniwersytet Pedagogiczny w Krakowie, Instytut Fizyki, Polska

<sup>2</sup> Doktor, Uniwersytet Pedagogiczny w Krakowie, Katedra Psychologii, Polska

<sup>3</sup> Doktor, Uniwersytet Pedagogiczny w Krakowie, Katedra Psychologii, Polska

#### Streszczenie

W artykule zawarto propozycję metod analizy stopnia złożoności obrazów dla celów badań eyetrackingowych. Problem niejednorodności parametrów fizycznych obrazu jest częstą przyczyną modyfikowania bodźców w celu ich ujednoczenia pod kątem tych parametrów. W przypadku badania odbioru dzieł sztuki taka operacja prowadzi do naruszenia struktury samego dzieła i rodzi problemy metodologiczne. Uwzględnienie współczynnika złożoności obrazu pozwala na porównanie danych eyetrackingowych uzyskanych z różnorodnych bodźców bez konieczności modyfikacji obrazu. Niniejszy artykuł opisuje sposób obliczania tego współczynnika w oparciu o gradient dla poszczególnych pikseli obrazu. Metoda ta może mieć zastosowanie w szeregu badań eyetrackingowych z wykorzystaniem złożonych bodźców.

**Słowa kluczowe:** *eyetracking*, stopień złożoności obrazu

#### Abstract

The article describes a method of analyzing the degree of complexity of images. Modification of images is often necessary when they differ with respect to statistical image properties such as complexity for instance. However, when it comes to studies on perception of art, such intervention might violate the actual structure of an artwork leading to severe methodological issues. We propose a different approach i.e. application of a coefficient aimed to control for unwanted effects of image complexity on eye movements. Present article describes the method of complexity coefficient calculation based on gradient counting for each pixel of an image. This approach might be useful in many studies that employ eyetracking to complex stimuli.

**Keywords:** *eyetracking*, image complexity, gradient function

---

Dynamiczny rozwój badań eyetrackingowych sprawia, że podejmowane są liczne próby stosowania tej metodologii w bardzo szerokim zakresie, zarówno w obszarach badań naukowych, jak i empirycznej optymalizacji, np. wyglądu

stron internetowych. Ze względu na ogromną różnorodność obiektów będących przedmiotem badań niezbędne jest opisywanie stopnia ich percepcyjnej złożoności. Wpływ czynników oddolnych, takich jak złożoność, na ruchy oczu czy odbiór subiektywny jest dobrze udokumentowany w literaturze (Fuchs, Ansorge, Redies, Leder, 2011; Jacobsen, Hofel, 2002). Wzrost złożoności obrazu koreluje zazwyczaj z długością oraz liczbą fiksacji, sugerując wzrost obciążenia poznawczego (np. Nuthmann, 2017). Na poziomie subiektywnym z kolei widzowie preferują zwykle obrazy o średnim poziomie złożoności (Berlyne, 1970). Zatem problem złożoności obrazu ma istotny wpływ na uzyskane wyniki badań.

W przypadku percepcji sztuki obrazy często różnią się złożonością, która z kolei wpływa zarówno na fizjologiczne, jak i subiektywne reakcje widzów. W konsekwencji niekontrolowana utrudnia interpretację wpływu głównych zmiennych, takich jak treść czy styl. W takiej sytuacji badacze często modyfikują oryginalne dzieła sztuki, mając na celu ujednoczenie obrazów pod względem złożoności. Zabieg ten rozwiązuje jeden problem, lecz niestety generuje inny związany z trafnością tak prowadzonych badań. Istotne staje się wówczas pytanie, czy faktycznie badanie dotyczy jeszcze sztuki, czy też po prostu bodźców wzrokowych. Dla obiektów tego typu wydaje się zasadne zastosowanie innych metod rozwiązania tego dylematu. Jedną z nich może być obliczenie współczynnika złożoności dla każdego obrazu, a następnie uwzględnienie go w modelach statystycznych.

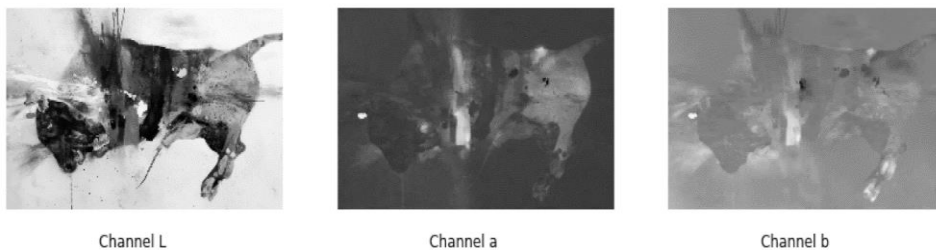


Rysunek 1. Tauromachie 002 (część 1), Massimiliano Carollo, 2016.

Źródło: przedruk za zgodą autora.

Procedurę obliczania współczynnika zilustrowano, posługując się obrazem przedstawionym na rysunku 1. Ponieważ w badaniach eyetrackingowych najczęściej wykorzystywane są obrazy cyfrowe, dlatego niezbędne jest dostosowanie ich wielkości, rozdzielczości i zastosowanie jednolitej specyfikacji definiującej luminancję oraz przestrzeń barw wyświetlanych obrazów. Do tego

celu wybrano przestrzeń barw określaną jako LAB (Braun, Amirshahi, Denzler, Redies, 2013). Jest to matematyczny opis trzech parametrów: luminancji  $L$ , kanałów  $a$  i  $b$  odpowiednio dla barw zieleni–czerwień, niebieski–żółty. Istotną zaletą jest fakt, iż przestrzeń kolorów LAB przekracza gamę kolorów modeli RGB i CMYK. Jednym z najważniejszych atrybutów modelu LAB jest niezależność od urządzenia. Oznacza to, że kolory są definiowane niezależnie od ich natury lub urządzenia, na którym są wyświetlane. Poniżej przedstawiono obraz z rysunku 2 z rozbiciem na trzy składowe.



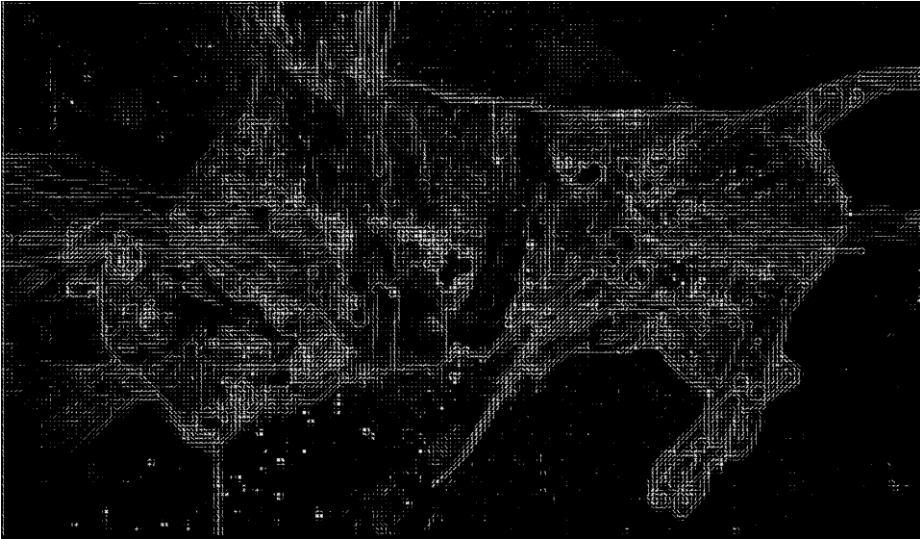
**Rysunek 2. Obraz Tauromachie 002 (część 1), Massimiliano Carollo, 2016 przedstawiony z rozbiciem na kanały a, b reprezentujące barwy oraz kanał luminancji**

Źródło: opracowanie własne.

Jednym z często stosowanych algorytmów w technikach komputerowej analizy obrazu jest histogram zorientowanych gradientów (HOG). Jest to deskryptor cech używany do wykrywania obiektów w komputerowej analizie, detekcji kształtów i przetwarzaniu obrazu. Technika deskryptorów HOG zlicza występowanie orientacji gradientu w zlokalizowanych częściach okna wykrywania obrazu lub regionu zainteresowania (ROI).

Implementacja algorytmu deskryptora HOG wygląda następująco. Dokonujemy podziału obrazu na małe połączone obszary zwane *komórkami*. Dla każdej komórki tworzymy histogramy gradientu dla poszczególnych kierunków lub orientacji krawędzi dla pikseli w komórce. Dokonujemy dyskretyzacji każdej komórki zgodnie z orientacją gradientu. Piksel każdej komórki ma przypisany gradient wraz z odpowiednią wagą oraz orientacją. Grupy sąsiednich komórek są uważane za regiony przestrzenne zwane *blokami*. Grupowanie komórek w blok jest podstawą do grupowania i normalizacji histogramów. Znormalizowana grupa histogramów reprezentuje histogram blokowy. Zbiór tych histogramów blokowych reprezentuje deskryptor. Według zaleceń (Dalal, 2005) optymalne wartości parametrów wynoszą odpowiednio: rozmiary okna detekcji 64 x 128, rozmiar komórki to 8 x 8, rozmiar bloku to 16 x 16 (2 x 2 komórki).

Na rysunku 3 przedstawiono implementację algorytmu HOG dla obrazu z rysunku 1 z zastosowaniem założonego progu binaryzacji o wartości 0,5.



Rysunek 3. HOG, efekt binaryzacji histogramu znormalizowanego dla założonego progu 0,5  
 Źródło: opracowanie własne.

Ze względu na fakt, iż algorytmy HOG przeznaczone są do innych celów, uznaliśmy, iż warto wprowadzić dodatkowy parametr opisujący badany obraz. Przyjęto za (Braun i in., 2013) parametr określający *stopień złożoności obrazu*. Współczynnik stopnia złożoności obrazu zdefiniowano następująco:

Zakładając, że funkcje:  $\nabla I_L(x, y)$ ,  $\nabla I_a(x, y)$ ,  $\nabla I_b(x, y)$  są gradientem, obliczanym kolejno dla pikseli o współrzędnych  $(x, y)$ , gdzie:  $0 < x < N$ ,  $0 < y < M$ , analizowanego obrazu dla poszczególnych kanałów luminancji  $L$  oraz barw  $a, b$ , to wybieramy wartość maksymalną gradientu dla danego punktu spośród wartości uzyskanych dla poszczególnych kierunków i kanałów:

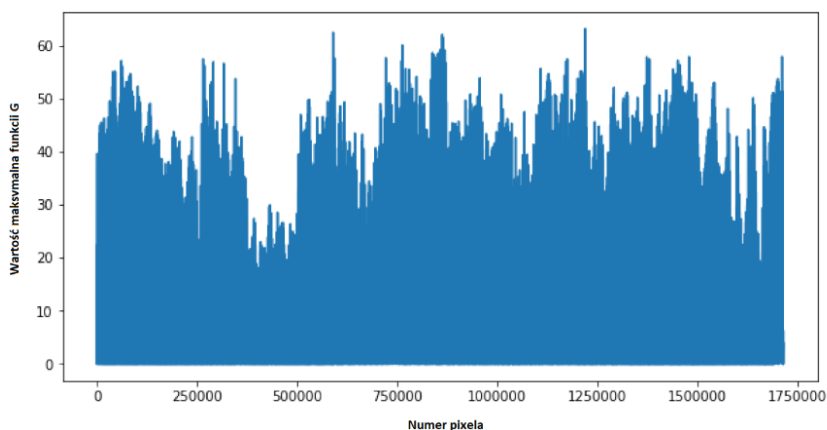
$$G_{\max}(x,y) = \max (|\nabla I_L(x, y)|, |\nabla I_a(x, y)|, |\nabla I_b(x, y)|).$$

Wówczas:

$$M_{Co} = (G_{\max}(x,y)) = \frac{1}{N \cdot M} \sum_{x,y} G_{\max}(x,y).$$

Wartości  $N, M$  to wysokość i szerokość, zatem iloczyn  $NM$  wyraża liczbę punktów.

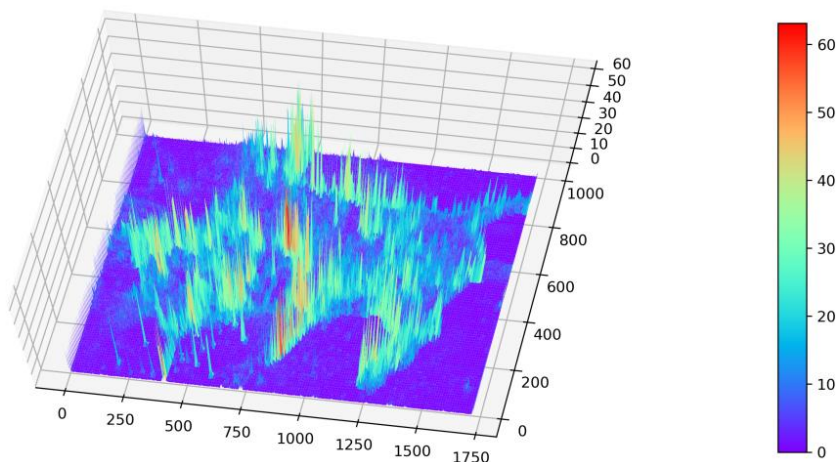
Dla całego zbioru uzyskanych wartości możemy policzyć wartość odchylenia standardowego. Wartość tę możemy interpretować jako opis stopnia niejednorodności obrazu. Dla obrazu z rysunku 1 wartość ta wynosi w przybliżeniu 4,32. Wartości maksymalne gradientu dla poszczególnych pikseli obrazu obliczone dla grafiki z rysunku 1 przedstawione w sposób linearny ilustruje kolejny histogram.



**Rysunek 4. Wartości maksimów gradientu dla obrazu z rysunku 21**

Źródło: opracowanie własne.

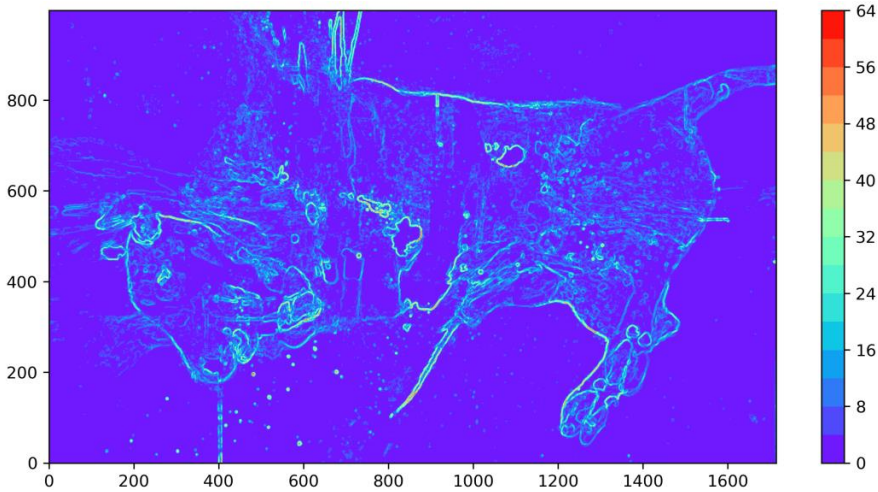
Ponieważ identyfikacja na obrazie maksimów gradientu dla poszczególnych punktów obrazu przedstawionych w sposób linearny jest dość trudna, można dokonać wizualizacji w trzech wymiarach. Kolejny rysunek przedstawia te same wartości parametru  $G_{\max}$  z zastosowaniem innych metod wizualizacji. Osie X, Y to współrzędne piksela, oś Z to wartość maksimów gradientu.



**Rysunek 5. Wartości maksimów gradientu dla poszczególnych punktów obrazu z rysunku 1– wizualizacja 3D**

Źródło: opracowanie własne.

Na rysunku przedstawiono te same dane wizualizowane z wykorzystaniem heatmapy.



**Rysunek 6. Wartości maksimów gradientu dla poszczególnych punktów obrazu z rysunku 1 – heatmapa**

Źródło: opracowanie własne.

Tak obliczona złożoność – czy to w postaci graficznej, czy też w postaci konkretnego współczynnika – może następnie zostać wykorzystana w modelach statystycznych jako jeden z czynników prognostycznych poszczególnych parametrów okulograficznych. Dodatkowo współczynnik ten koreluje z szeregiem reakcji subiektywnych. W badaniu Lyssenko, Redies i Hayn-Leichsenring (2016) zaobserwowano, że współczynnik złożoności w oparciu o gradienty koreluje pozytywnie ze złożonością subiektywną ( $r = 0,68$ ;  $p < ,001$ ), poziomem nieustrukturyzowania obrazu ( $r = 0,65$ ;  $p < ,001$ ) czy jego barwnością ( $r = 0,2$ ;  $p < ,001$ ). W zrealizowanym przez nas badaniu dokonano analizy korelacji pomiędzy współczynnikiem złożoności a subiektywnymi szacowaniami złożoności. W badaniu osoby ( $N = 94$ ) oglądały swobodnie 20 obrazów współczesnych. Po obejrzeniu obrazów miały oszacować subiektywnie spostrzeganą złożoność obrazu wyrażoną w skali od 1 do 5, gdzie 5 oznaczało największy poziom złożoności. Korelacja pomiędzy współczynnikiem złożoności obliczonym zgodnie z procedurą zaproponowaną powyżej a subiektywnie szacowaną złożonością obrazu wynosiła 0,18 ( $p < ,05$ ). Zatem mimo iż istotna statystycznie, była dużo niższa niż we wspomnianym wcześniej eksperymencie. Różnica pomiędzy tymi dwiema korelacjami może wynikać z różnic w użytych bodźcach. W eksperymencie Lyssenko, Redies i Hayn-Leichsenring (2016) użyto wyłącznie obrazów abstrakcyjnych, zatem subiektywne szacowanie złożoności opierało się głównie na wartościach formalnych. W naszym badaniu użyto obrazów figuratywnych, a więc szacowanie złożoności mogło dotyczyć również złożoności semantycznej.



Kolejnym celem badania była próba weryfikacji celowości obliczania stopnia złożoności obrazu do przewidywania obszarów, dla których liczba i czasy fiksacji będą największe. Na rysunku 7 przedstawiono uzyskane w badaniu złożenie zapisów *scanpath* dla wszystkich jego uczestników w czasie pierwszych 10 sekund analizy – czasie, w którym dominuje całościowe skanowanie obrazu oraz procesy oddolne (Locher, Krupinski, Mello-Thomas, 2007; Henderson, Weeks, Hollingworth; Itti, Koch, 2000). Potwierdzają to liczne stosunkowo krótkie fiksacje i duże amplitudy ruchów sakkadycznych.



Rysunek 7. ScanPath dla pierwszych 10 sekund analizy – wszyscy badani

Źródło: opracowanie własne.

## Podsumowanie

Współczynnik złożoności wydaje się istotnym elementem, który należy kontrolować. Może on mieć wpływ zarówno na subiektywne szacowania, jak i ruchy oczu osób badanych. Uwzględnienie go w modelu statystycznym pozwala na korygowanie wpływu pozostałych parametrów na interesujące badacza zmienne, takie jak liczba czy długość fiksacji bądź preferencje estetyczne. Ponadto heatmapy zbudowane w oparciu o opisany algorytm mogą być przydatne do przewidywania lokalizacji fiksacji, zwłaszcza w sytuacji bodźców, w których jednoznaczne wyznaczenie obszarów zainteresowania może być problematyczne. Dalsze i bardziej szczegółowe analizy pozyskanych danych eksperymentalnych będą przedmiotem kolejnych publikacji.

## Literatura

- Berlyne, D.E. (1970). Novelty, Complexity and Hedonic Value. *Perception and Psychophysics*, 8, 279–286.
- Braun J., Amirshahi, S.A., Denzler, J., Redies, C. (2013). Statistical Image Properties of Print Advertisements, Visual Artworks and Images of Architecture. *Frontiers in Psychology*, 4, 1–15. DOI: 10.3389/fpsyg.2013.00808.
- Dalal, N., Triggs, B. (2005). *Histograms of Oriented Gradients for Human Detection*. INRIA.
- Fuchs, I., Ansoorge, U., Redies, C., Leder, H. (2011). Saliency in Paintings: Bottom-up Influences on Eye Fixations. *Cognitive Computation*, 3(1), 25–36.
- Henderson, J.M., Weeks Jr, P.A., Hollingworth, A. (1999). The Effects of Semantic Consistency on Eye Movements during Complex Scene Viewing. *Journal of Experimental Psychology: Human Perception and Performance*, 25(1), 210–228.
- Itti, L., Koch, C. (2000). A Saliency-based Search Mechanism for Overt and Covert Shifts of Visual Attention. *Vision Research*, 40(10–12), 1489–1506.
- Jacobsen, T., Höfel, L.E.A. (2002). Aesthetic Judgments of Novel Graphic Patterns: Analyses of Individual Judgments. *Perceptual and Motor Skills*, 95(3), 755–766.
- Jain, A.K. (1989). *Fundamentals of Digital Image Processing*. New Jersey, United States of America: Prentice Hall.
- Locher, P., Krupinski, E.A., Mello-Thoms, C., Nodine, C.F. (2007). Visual Interest in Pictorial Art during an Aesthetic Experience. *Spatial Vision*, 21(1), 55–77.
- Lyssenko, N., Redies, C., Hayn-Leichsenring, G.U. (2016). Evaluating Abstract Art: Relation between Term Usage, Subjective Ratings, Image Properties and Personality Traits. *Frontiers in Psychology*, 7(973), 1–9.
- Nuthmann, A. (2017). Fixation Durations in Scene Viewing: Modeling the Effects of Local Image Features, Oculomotor Parameters, and Task. *Psychonomic Bulletin & Review*, 24(2), 370–392. [http://doi 10.3758/s13423-016-1124-4](http://doi.10.3758/s13423-016-1124-4).