

---

# Fotografia w bibliotece i w sieci: dwa podejścia do reprezentacji i wyszukiwania

Kamil Stępień

*Institut Informacji Naukowej i Bibliotekoznawstwa  
Uniwersytet Marii Curie-Skłodowskiej w Lublinie*

Jadwiga Woźniak-Kasperek

*Katedra Bibliografii i Dokumentacji, Wydział Dziennikarstwa, Informacji i Bibliologii  
Uniwersytet Warszawski*

---

## Abstrakt

**Cel/Teza:** Celem jest przedstawienie dwóch podejść do reprezentowania i wyszukiwania fotografii – tradycyjnego, metadaneowego, specyficznego dla bibliotek z automatycznym, sieciowym. Rozważania zogniskowane są na możliwości ich wzajemnego uzupełniania się w celu stworzenia nowej jakości informacyjno-wyszukiawczej w wielokanałowej komunikacji międzyludzkiej.

**Koncepcja/Metody badań:** Wykorzystano metodę analizy i krytyki piśmiennictwa.

**Wyniki i wnioski:** Optymalnym rozwiązaniem problemów reprezentacji i wyszukiwania fotografii byłby dziś system hybrydowy łączący opis tworzony przez profesjonalistę z treściami generowanymi oddolnie przez internautów oraz z rezultatami komputerowej analizy obrazu. Opis obiektu wizualnego tworzony równoległe na różnych płaszczyznach pozwoliłby na zmniejszenie ryzyka powstawania opisów błędnych, niezrozumiałych i niekompletnych oraz na wzrost możliwości wyszukiwania obiektów wizualnych z uwzględnieniem zindywidualizowanych potrzeb użytkowników.

**Oryginalność/Wartość poznawcza:** Artykuł uzupełnia niedostatecznie reprezentowaną w polskiej literaturze bibliologiczno-informatologicznej problematykę komputerowej analizy obrazu, zwłaszcza fotografii, i jej użyteczności informacyjno-wyszukiawczej.

## Słowa kluczowe

Biblioteka. Fotografia. Komputerowa analiza obrazu. Reprezentacja. Wyszukiwanie.

*Otrzymany: 14 lutego 2019. Zrecenzowany: 28 lutego 2019. Zaakceptowany: 14 marca 2019.*

---

## 1. Wstęp

Środowisko wyszukiwawcze Internetu w znacznej części stanowią obrazy, treści wizualne zestawione na ogół z mniejszą lub większą porcją tekstu. „Zestawienie tekstu i ilustracji składa się na przekaz o «podwójnej konceptualizacji»” (Tabakowska, 2006, 10). Pewną część ikonosfery sieciowej stanowią cyfrowe fotografie, w tym również cyfrowe kopie fotografii przechowywanych w instytucjach takich jak biblioteki, wraz z ich metadanymi. W artykule chcemy porównać na przykładzie fotografii dwa „paradygmaty” reprezentowania i wyszukiwania obrazów – biblioteczny, „tradycyjny”<sup>1</sup>, bazujący na metadanych z sieciowym,

---

<sup>1</sup> W artykule ograniczamy perspektywę do sfery bibliotek, zdając sobie sprawę, że nie tylko one dysponują zasobami fotografii oraz że podobnie postępuje się z fotografiami w innych instytucjach należących

możliwym do stosowania także w bibliotekach, choć na pewno nie bez problemów i łamania pewnych schematów. Zawarte w artykule treści dotyczące komputerowego rozumienia obrazu i wyszukiwania obrazem nie mają charakteru poradnika implementacyjnego, ale tworzą podstawową ramę konceptualną, z której można wywieść wnioski praktyczne.

Odwołując się do języka specjalistycznego bibliotekarzy i pracowników informacji, można powiedzieć, że w artykule będziemy skupiać uwagę na determinantach możliwości wyszukiwania (ang. *retrieval*) fotografii oznaczającego „kolejne, indywidualne realizacje wyszukiwania (...), obejmujące ciąg interakcji jednostki poszukującej informacji z systemem informacyjnym, zarówno na poziomie manipulowania systemem, jak i na poziomie intelektualnym, tj. tworzenia instrukcji wyszukiwawczej, ustalania kryteriów selekcji wyszukanej informacji oraz dotyczącym oceny trafności wyszukanej informacji” (Wilson, 2000, 49)<sup>2</sup>, a także „odnajdywania” (ang. *seeking*) będącego procesem mniej sformalizowanym i ustrukturyzowanym, ale coraz bardziej spersonalizowanym i złożonym, angażującym technikę komputerową. W praktyce sieciowej szukający płynnie przechodzi od pytania przez przeglądanie, filtrowanie, po wyszukiwanie, by wrócić do pytania itd. Wyszukiwanie stwarza bowiem szansę na lepsze zrozumienie problemu i ewentualne powtórzenie czynności. Można zaobserwować dynamiczny rozwój łączenia sposobów wizualizacji treści z samym wyszukiwaniem, co daje pożądane rezultaty nie tylko w postaci bogatego zestawu hiperłączy, ale także połączeń z innego rodzaju treściami, takimi jak: multimedia, obrazy, tłumaczenia czy dane geograficzne. Jesteśmy przekonani, że metadane słowne (opisy, reprezentacje przedmiotowe, rzeczowe) są nadal niezbędnym elementem środowiska wyszukiwawczego, ale im poświęcono relatywnie dużo publikacji, na które zwracamy uwagę zainteresowanych czytelników (Baca 2002; Baca 2006; Beghtol 1986; Clough, Ireson, Marlow, 2009; Jörgensen, 2003; Jörgensen et al., 2001; Jörgensen, 1999; Jörgensen, 1998; Rygiel, 2012; Shatford-Layne, 1994; Shatford, 1986; Svenonius, 1994; Włodarczyk & Woźniak-Kasperek, 2017).

## 2. Fotografia w bibliotece – poza tradycyjne kategorie

Dokumenty ikonograficzne, a wśród nich fotografie, zgromadzone w bibliotekach, w muzeach, w archiwach i innych instytucjach kultury i informacji są opracowywane i przygotowywane do wyszukiwania zazwyczaj w podobny, „konwencjonalny” sposób, właściwy dla danego środowiska. Zasoby te stanowią nadal nie w pełni doceniane źródło (Talbierska 2009), mimo że przekaz za pomocą kodu wizualnego pozwala na transmisję tych treści, które są niemożliwe lub trudne do wyrażenia w inny sposób, co ma określoną wartość poznawczą, a mieści się w sferze zainteresowania różnych dziedzin wiedzy i dyscyplin naukowych, m.in. historii sztuki, socjologii, etnologii, architektury czy urbanistyki. Specyficzne dla bibliotek podejście do opracowania fotografii sprawia, że ich wyszukiwanie może być realizowane poprzez odwołanie jedynie do kategorii przyjętych na etapie tworzenia reprezentacji.

---

do klasy GLAM (Galleries, Libraries, Archives, Museums). Przyjeliśmy również założenie, że sama zmiana kodu z analogowego na cyfrowy nie zmienia statusu fotografii, inaczej mówiąc, zdigitalizowanie fotografii i umieszczenie jej w bibliotece cyfrowej nie czyni z niej automatycznie fotografii sieciowej.

<sup>2</sup> Wszystkie cytaty z prac nieopublikowanych dotychczas w języku polskim są tłumaczeniami autorów niniejszego artykułu [red.].

Analizę i opracowanie rzeczowe fotografii w bibliotece można zrealizować na kilka sposobów. Szczegółowy opis procedury analizy i tworzenia opisu zawiera m.in. publikacja B. Włodarczyka i J. Woźniak-Kasperek (Włodarczyk & Woźniak-Kasperek, 2017). W rezultacie powstają opisy zawierające odpowiedzi na pytania o: przedmiot główny widoczny na fotografii, przedmioty równorzędne i poboczne, własności przedmiotów, w tym lokalizację przestrzenną (geograficzną) i czasową, ich ujęcie (punkt widzenia), adresatów, formę (np. fotografia czarno-biała polska), czy postać fizyczną (np. odbitka srebrowo-żelatynowa). W przypadku zasobów o dużym wolumenie podejście takie może powodować generowanie tak licznych odpowiedzi, że wyszukiwanie właściwie należy uznać za nieudane. Użytkownik nie jest bowiem w stanie zapoznać się z całością otrzymanych wyników, a nie ma do dyspozycji narzędzi umożliwiających ich ograniczenie. Z różnych powodów (brak możliwości głębszej, poprawnej i kompletnej interpretacji poszczególnych warstw treści fotografii, wiedzy o nośniku i technikach utrwalenia, niedostatków stosowanych słowników kontrolowanych lub niekontrolowanych itd.) nie jest bowiem możliwa rafinacja odpowiedzi. Z pomocą mogłyby tu przyjść techniki automatycznego rozpoznawania i wyszukiwania obrazu obrazem (wyszukiwanie odwrotne), o czym będzie mowa w dalszych częściach artykułu.

Digitalizacja i prezentowanie cyfrowych kopii i cyfrowych oryginałów kolekcji muzealnych, bibliotecznych czy archiwalnych stały się w ostatnich latach powszechną praktyką. Obecność dokumentów ikonograficznych w środowisku sieciowym sprawia, że z jednej strony, uaktualniane są w duchu nowych możliwości i potrzeb odbiorców praktyki transformacji danych i prezentowania informacji o nich oraz sposób ich utrwalenia, z drugiej zaś, wprawdzie jeszcze nieśmiało, ale instytucje te zaczynają czerpać z oferty zaawansowanych technologii cyfrowych. Cyfryzacja zasobów oraz umieszczenie ich w internecie miały wpływ na zmiany w strukturze i prezentacji danych opisu. Nadal jednym z nierozwiązanych problemów jest jednak problem takiego języka opisu rzeczowego, systemu organizacji wiedzy, ontologii itp., które odpowiadałyby potrzebom informacyjnym różnych odbiorców (indywidualnych, ale też instytucjonalnych) oraz specyfice obrazu. Brakuje systemowego podejścia do opisu przedstawień wizualnych, w tym tu nas interesujących fotografii, które sprzyjałoby konsekwencji w tworzeniu poprawnych i kompletnych reprezentacji fotografii, a co za tym idzie, eliminowaniu błędów i trudności wyszukiwawczych. Niezbędna jest rewizja dotychczasowych rozwiązań, wsparta pogłębioną refleksją nad sposobami tworzenia i zarządzania metadanymi opisowymi oraz wzbogacenie repertuaru tradycyjnych metod i środków o te, które są dostępne dzięki technologiom przetwarzania obrazów cyfrowych.

Zbiory ikonograficzne są cennym, choć nie priorytetowym, typem materiałów gromadzonych w bibliotekach, gdzie najczęściej są przechowywane, opracowywane i udostępniane w osobnych działach, jako specjalistyczne kolekcje na przykład o charakterze artystycznym (grafika, rysunek, fotografia). Mogą też wchodzić w skład większych kolekcji, np. dokumentów życia społecznego. „Jeszcze do niedawna zbiory fotograficzne były często lekceważone. Fotografie katalogowano zbiorczo, bardziej jako kolekcje, czasem w ogóle nie opisywano” (Miller & Wombard, 2009, 202). Przechowywano je luzem, w pudełkach, teczkach, bez szczegółowego komentarza. Negatywy fotograficzne – w kopertach czy tekturowych pudełkach „przechowywano [...] razem z dokumentami papierowymi, a jeszcze w latach 80. XX w. twierdzono czasem, że w ogóle nie są archiwaliami” (Miller & Wombard, 2009, 222). Ich tradycyjny opis, dotyczący zawartości, treści przedstawień,

zawsze stwarzał w środowisku bibliotecznym duże problemy, ale też trzeba przyznać, że nie poświęcono im tyle uwagi co dokumentom piśmienniczym. Fotografie w zbiorach bibliotek, archiwów i innych instytucji informacji i kultury są opisywane i porządkowane stosownie do przyjętych zwyczajów i standardów, co nie oznacza, że ich odszukanie jest zawsze łatwe i skuteczne. Przyczyn tego stanu rzeczy jest kilka. Jedną, jak wcześniej wspomnieliśmy, jest preferowanie dokumentów piśmienniczych kosztem dokumentów innych typów, skutkiem czego mechanizmy i narzędzia indeksowania i wyszukiwania dokumentów piśmienniczych są lepiej rozpoznane, bogatsze i skuteczniejsze. Trudności nastręcza sama praca z dokumentem wizualnym, poznawanie go, rozpoznawanie, interpretowanie. Niejednokrotnie brakuje danych tekstowych, dokumentów towarzyszących, zapisków na odwrociach itp., które pomogłyby odpowiedzieć na pytania kluczowe dla opracowania zasobu, brakuje także klarownych instrukcji regulujących sposób tworzenia opisów i generowania metadanych.

Do najważniejszych zadań bibliotekarza, który ma do opracowania fotografię bez żadnego tekstu, należą: identyfikacja i datowanie, określenie miejsca, przedmiotów widocznych na fotografii, zidentyfikowanie portretowanej osoby, autora fotografii itp. Wymaga to szerokiej wiedzy. Aby mieć pewność, że przypuszczenia odnośnie do datowania są słuszne, należy znać techniki i technologie wykonywania fotografii, szczególnie jeżeli mamy do czynienia z fotografią starą, przydatna jest także znajomość historii sztuki, m.in. stylów i panującej wówczas mody, niezbędna jest umiejętność odróżnienia „czystej” fotografii od jej drukowanej wersji itd. Opracowanie fotografii oprócz analizy i nazwania elementów przedstawienia, które się nań składają (ang. *ofness*), oraz „cech fizycznych dokumentu, (...), czym jest zasób, dzieło, jaka jest jego forma” (Rygiel, 2012, 78) (ang. *isness*, co jest bliskie kategorii gatunek/forma), może zawierać odpowiedź na pytanie o symbolikę, o to czego dotyczy, o czym jest, jakie niesie przesłanie itp. (w literaturze anglojęzycznej ta ścieżka opracowania nosi nazwę *aboutness* i eksplorujemy ją, opracowując wiele innych typów dokumentów). „Pojęcia *aboutness* i *ofness* stosowane w opisie obrazów oparte są na teorii Panofsky’ego, dotyczącej trzech warstw znaczeniowych dzieła sztuki oraz koncepcji OF i ABOUT Shatford” (Rygiel, 2012, 78–79). Najniższy poziom, pre-ikonograficzny, idąc śladem Panofsky’ego, lub *ofness* według Shatford tworzą zidentyfikowane i nazwane elementy opracowywanego przedstawienia. Poziom ikonograficzny wymaga interpretacji wyodrębnionych elementów *ofness* i ich konfiguracji z uwzględnieniem wiedzy przedmiotowej oraz konwencji i kodów artystycznych, społecznych, kulturowych. Trzeci poziom, ikonologiczny, wymaga rozpoznania warstwy symbolicznej przedstawienia, zrekonstruowania programu ikonograficznego, co na ogół dla osoby niemającej odpowiedniej wiedzy i doświadczenia może być zadaniem niewykonalnym. Podejście ikonograficzne zakłada czytanie sztuki, w tym przypadku fotografii, czyli interpretowanie wizualnej reprezentacji, przez umieszczanie jej elementów w tradycji, która daje im znaczenie inne niż to, które sugeruje ich wizualny wygląd (Bał, 1991, 177). Zakładając, że etap pracy z dokumentem zostanie zrealizowany z sukcesem, kolejne czynności związane z wyrażeniem treści z wykorzystaniem stosowanego w systemie słownika nie różnią się od tych podejmowanych przy opracowaniu dokumentów piśmienniczych, zakładając, że słownik jest dostatecznie nasycony relewantnym słownictwem. W dalszej części artykułu wrócimy do tych zagadnień, jako że automatyzacja wprawdzie nie gwarantuje bezwzględnie lepszych rezultatów, ale pozwala je wzbogacić i przyspieszyć, co też nie jest bez znaczenia.

Wspomniany wcześniej amerykański historyk sztuki Erwin Panofsky wydzielił trzy poziomy analizy dzieła sztuki: pre-ikonograficzny, ikonograficzny i ikonologiczny. Podział ten można zastosować podczas analizy przedmiotowej. Bibliotekarz opracowujący zdjęcie w modelu tradycyjnym skupia się na dwóch pierwszych poziomach, pomijając znaczenie symboliczne przedstawienia. Uwzględnienie poziomu pre-ikonograficznego pozwala opisać kategorie obiektów przedstawionych na fotografii. Jednak ich poprawne nazwanie jest możliwe dopiero po identyfikacji indywidualnych osób, ich kategorii, obiektów itd. Ideałem byłoby określenie znaczenia symbolicznego przedstawienia, o ile w konkretnym przypadku można o nim mówić, a co w bibliotekach ma miejsce niezwykle rzadko. Weźmy jako przykład poniższą fotografię (Rys. 1) z biblioteki cyfrowej Polona i jej opis, w którym znajdziemy pewne elementy z kategorii *ofness* i *isness*. Autor opracowania nadał fotografii tytuł (Wnętrze laboratorium szpitala wojennego w Kielcach), zrezygnował z próby jej interpretacji symbolicznej, a nawet z głębszej analizy warstwy *ofness*. Charakterystyka wyszukiwawcza dokumentu sporządzona w bibliotece cyfrowej Polona składa się z następujących haseł – słów kluczowych: *Kielce (woj. świętokrzyskie); Lekarze; Szpitale wojskowe*, datacji powstania fotografii: *1914–1918*, określenia gatunku: *Odbitka na papierze srebrowo-żelatynowym; Fotografia polska* oraz opisu fizycznego, tj. *1 fot. : odb. na pap. srebrowo-żelatynowym; 8,9x13,9 cm*. W rezultacie również wyszukiwanie jest możliwe jedynie na poziomie tak wyodrębnionych kategorii. Poniżej przedstawiamy tagi przypisane tej fotografii przez grupę internautów, studentów kierunku architektura informacji.



Rys. 1. Wnętrze laboratorium szpitala wojennego w Kielcach [1914–1918].

Źródło: Polska Biblioteka Cyfrowa Polona [online], <https://polona.pl/item/wnetrze-laboratorium-szpitala-wojennego-w-kielcach,N7Tc4NjkkOA/0>, [30.01.2019]

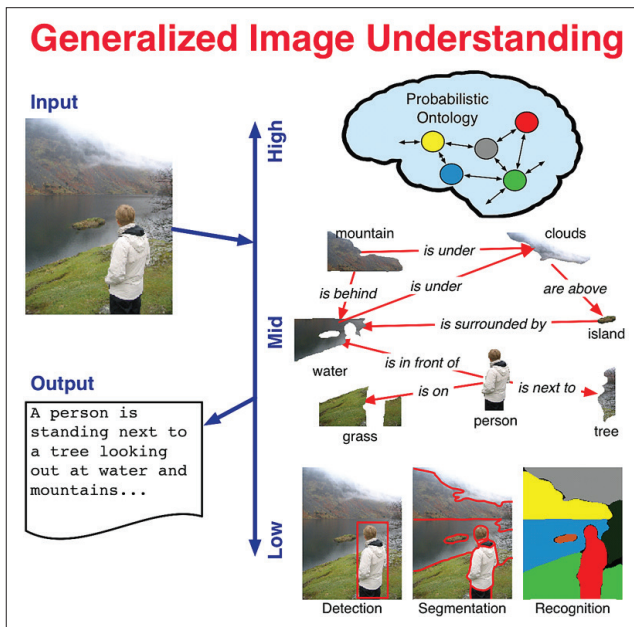
Tab. 1. Zestawienie tagów użytkowników z opisem w bibliotece Polona dla fotografii  
*Wnętrze laboratorium szpitala wojennego w Kielcach [1914–1918]*

Uporządkowane alfabetycznie tagi, nadane przez internautów	Opis w Bibliotece Cyfrowej Polona
<p>ampułki, apteczka, apteka, archiwum, badania, badanie, badanie krwi, badanie lekarskie, białe fartuchy, biały fartuch, biały kitel, biel, biologia, blat, butelka, chemia, chemicy, chemikalia, chirurdzy, choroba, choroby, chory, ciekawość, czarno białe, czarno-biała, czarno-białe, czerwony krzyż, czysto, dawna medycyna, dawne czasy, dawne leczenie, dawno temu, dezynfekcja, doktor, doktorzy, doświadczenie, eksperyment, eksperymenty, faceci w bieli, farmaceuci, fartuch, fartuchy, fiolka, fiołki, fotografia, gabinet, gabinet lekarski, historia, hospital, igła, kitel, kitel lekarski, kitle, kitle lekarskie, krew, kwarantanna, laboratorium, laboratorium medyczne, lecznictwo, lek, lekarstwa, lekarstwo, lekarz, lekarze, lekcja medycyny, leki, ludzie, medycy, medycyna, mężczyzna, mężczyźni, młodzi ludzie, mundur, mundur lekarski, nauka, naukowcy, odbitka, odczynniki, oddawanie krwi, odwaga, opatrunek, operacja, organizm, pacjenci, pacjent, pielęgniarz, pielęgniarze, pierwsza pomoc, pipeta, pipety, pobieranie, pobieranie krwi, pomieszczenie, pomoc, półka, praca, pracownia, pracownicy, praktyki, próbówka, próbówki, próbki laboratoryjne, przyrządy medyczne, ramka, retro, roztwór, sepia, skupienie, spirytus, stare kitle lekarskie, stare narzędzia, stare zdjęcie, staroświeckie, sterylność, stetoskop, stół, strzykawka, strzykawki, studenci, studenci medycyny, substancja, szczepienie, szczepionka, szczepionki, szkolenie, szkoła, szpital, ściana, termometr, testy, ukłucie, uniform, uzdrowienie, vintage, wkłucie, wojsko, zabieg, zafascynowanie, zajęcia, zastrzyk, zdjęcie, zdrowie</p>	<p>adres wydawniczy: [1914–1918]  data powstania dokumentu: [1914–1918]  opis fizyczny: 1 fot.: odb. na pap. srebrowo-żelatynowym ; 8,9x13,9 cm  gatunek: odbitka na papierze srebrowo-żelatynowym  fotografia polska  forma i typ: Fotografie  słowa kluczowe: Szpitale wojskowe Lekarze Kielce (woj. świętokrzyskie)  uwagi: Tyt. nadany przez katalogującego.  sygnatura: E.82315  prawa do utworu: Domena Publiczna. Wolno zwielokrotnić, zmieniać i rozpowszechnić oraz wykonywać utwór, nawet w celach komercyjnych, bez konieczności pytania o zgodę. Wykorzystując utwór należy pamiętać o poszanowaniu autorskich praw osobistych Twórcy.  źródło: Biblioteka Narodowa  rekord bibliograficzny: zobacz w katalogu  uid repozytorium: 5786918</p>

### 3. Fotografia w sieci – rozpoznawanie obrazu, kształt, barwa, jasność

Odczytywanie każdego tekstu, także wizualnego, jest wielopoziomowe, złożone, a gdy czyni to człowiek, dodatkowo mniej lub bardziej subiektywne, interpretatywne, uczestniczące. Inaczej dzieje się w przypadku technik komputerowego rozpoznawania obrazów,

dla których trudnością jest nie identyfikacja kształtów, barw, tekstury itp. czy nazwanie rozpoznanych elementów, ale odkrycie powiązań (relacji, zależności) pomiędzy nimi, zinterpretowanie kontekstu różnych poziomów. Komputer rozpoznaje obiekty w sposób szczegółowy, ale nie zawsze poprawnie je identyfikuje i kategoryzuje (uogólnia). Problemem jest efektywność semantyki wizualnej, łączącej wiedzę językową i pozajęzykową z rozpoznawaniem wzorców przez komputer. Semantyka wizualna ma wesprzeć człowieka w opisywaniu i wyszukiwaniu obrazów, ale na tym etapie jeszcze nie zastąpi ludzkiej inteligencji, choć w pewnych sytuacjach opis wykonany przez człowieka jest nieosiągalnym luksusem dodatkowo naznaczonym zawieszeniem w sieciach znaczeń, które osoba ta sama utkała (Geertz, 2005, 19). Trudności, zarówno dla komputera, jak i dla człowieka, pojawiają się przy podejmowaniu prób interpretacji przedstawienia. Czy osoba na fotografii poniżej (Rys. 2) jest turystą/turystką, czy może właścicielką/właścicielem schroniska z niepokojem spoglądającym na spowite mgłą góry, a może jeszcze kimś innym? Czy jest to zwykłe zdjęcie, czy może zapis chwili przed lub po bardzo ważnym wydarzeniu? I tak dalej.



Rys. 2. Zrozumieć obraz. Źródło: Corso, 2014

Większość operacji w ramach widzenia komputerowego dotyczy rozpoznawania wzorców (ang. *pattern recognition*), czyli identyfikacji struktur w obrębie obrazu. Komputerową analizę obrazu można zatem sprowadzić do takich czynności jak:

- analiza ogólnych cech obrazu bądź jego fragmentu – analiza cech ilościowych (próbkowanie, pomiary kolorów);
- ekstrakcja cech – odseparowanie od siebie kształtów i obszarów w warstwie wizualnej dzięki zbadaniu krawędzi obiektów i ich struktury geometrycznej;
- segmentacja – podział obszaru wizualnego na spójne obszary o podobnych cechach, takich jak kolor, kształt;

- analiza tekstur – wydzielenie z obrazu charakterystycznych i rozpoznawalnych struktur takich jak np. struktura drewna, bruku;
- wydzielenie i śledzenie obiektów semantycznych, tj. takich, które posiadają złożone znaczenia, w których są wykazywane relacje zależności, np. dom, człowiek, samochód;
- śledzenie zmiany położenia tych obiektów ruchem własnym (przemieszczanie się) – przydatne w przypadku badania sekwencji obrazów statycznych, bądź cyfrowych obrazów animowanych typu GIF;
- reprezentacja obrazów i opis kształtów – sporządzanie syntetycznych opisów obszarów wydzielonych podczas procesu segmentacji: opisy kształtów, obiektów płaskich i przestrzennych;
- analiza sceny i zdarzeń – określanie zależności czasowo-przestrzennych między obiektami, analiza ruchu obiektów, klasyfikowanie ujęć i zdarzeń (Domański, 2010, 24).

Współczesne zdjęcia cyfrowe składają się z kilkudziesięciu milionów pikseli. „Porównywanie wszystkich, piksel po pikselu nawet przy obecnych mocach obliczeniowych komputerów jest zadaniem nierealnym” (Staszczak, 2010). Dlatego stosuje się rozwiązania upraszczające, polegające na określaniu tzw. wektorów cech. „Podczas wyszukiwania podobnego zdjęcia proces ten rozpoczyna się od wygenerowania wektora cech na podstawie wybranego identyfikatora, porównania go z innymi wektorami cech obrazów umieszczonych w bazie, a następnie pogrupowania i zobrazowania wyników” (Staszczak, 2010). W praktyce proces wyszukiwania danych graficznych w zbiorach obrazów odbywa się przez odpowiednie indeksowanie wektorów. Indeks w tym przypadku jest cechą obiektu o relatywnie małym stopniu złożoności. Indeksy tworzone są wprost z danych obrazowych, przez wybranie ich określonych własności całościowych bądź fragmentarycznych przy użyciu funkcji matematycznych. „Bardzo często indeks zawiera informacje, które nie są wprost reprezentowane w danych” (Forczmański, 2005). Wprowadzono rozróżnienie indeksów na niski i wysokopoziomowe (opisowe). Pierwsze z nich dotyczą informacji graficznej o cechach percepcyjnych i fizycznych, takich jak kształt, kolor, tekstura itp. Indeksy wysokopoziomowe są z kolei hasłami kategoryzującymi, tworzonymi na podstawie zawartości i metadanych.

Segmentacja obrazu przez wydzielenie w nim charakterystycznych cech obiektów skutkuje powstawaniem deskryptorów punktowych, krawędziowych i liniowych. Najpopularniejszym sposobem reprezentacji obiektów na obrazie jest ich kształt: kontur i krawędź. Rozwiązanie to sprawdza się bardzo dobrze wtedy, „gdy rozpoznawaniu podlegają konkretne, izolowane obiekty, m.in. w analizie obrazów medycznych (np. kształty komórek na obrazach mikroskopowych lub organy wewnętrzne ciała ludzkiego na obrazach tomograficznych) i optycznym rozpoznawaniu znaków (OCR)” (Forczmański, 2005). W tej sytuacji wszystkie inne cechy charakterystyczne są pomijane. Kontrast, jasność i barwa nie są tak stabilnymi cechami jak kształt. Kształt, którego odseparowane elementy przybierają postać binarną, składającą się z konturu i wnętrza są odporne na częściowe zasłonięcie obiektu, złej jakości odwzorowanie jego krawędzi czy szumy wynikające z niedoskonałości urządzeń rejestrujących obrazy. Najlepszym z rozwiązań jest „samouczenie się” algorytmów rozpoznających kształty, począwszy od tych najprostszych, płaskich figur typu elipsa, prostokąt, przez modele trójwymiarowe, a na odstępstwach i wariacjach bazujących na elementach skomplikowanych skończywszy. Rozwiązania tego typu są już częściowo wdrażane. Wszystko to dotyczy jednak warstwy logicznej obrazu, nie jego treści, interpretacji.



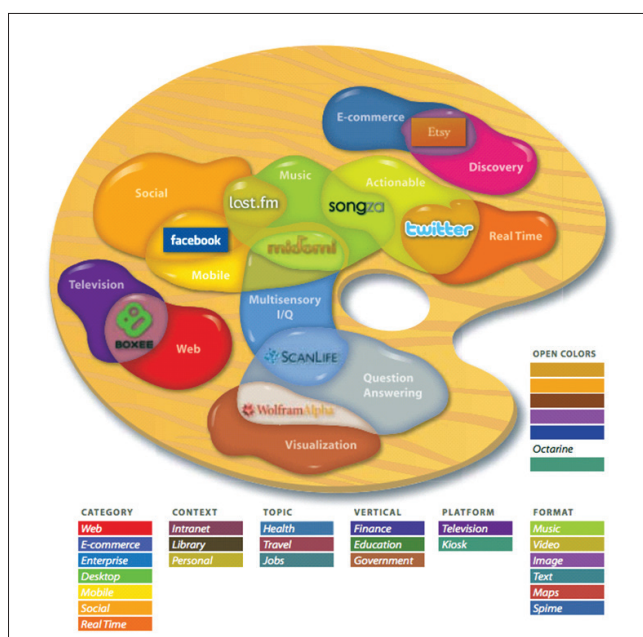
Deskrytory kształtu są wrażliwe na obroty, przesunięcia i zniekształcenia. Dodatkowo, mają niską złożoność obliczeniową. Tą metodą wyszukiwane są kształty obiektów, podobne lub takie same, jak założone na wejściu. Algorytm bada obraz na poziomie pikseli, wyodrębniając kontury przedmiotów za pomocą różnic w przejściach kolorystycznych (gradientów). Dzięki tej metodzie rozpoznawane obiekty charakteryzują się określonym kształtem, rodzajem tego kształtu, ale nie wielkością. „Zbieramy wszystkie właściwości, które oddają rzucające się w oczy szczegóły geometryczne obrazu” (Gevers & Smeulders, 2003). Za pomocą *shape searching* możliwe było stworzenie programów rozpoznających tekst ze skanowanych dokumentów (oprogramowanie typu OCR umożliwiające rozpoznawanie znaków w plikach graficznych). Oczywiście jest, że, aby rozpoznawanie miało określoną skuteczność, trzeba zdefiniować parametry, takie jak: alfabet znaków wejściowych do rozpoznawania, znaki nieliterowe, a także możliwość odróżnienia elementów w postaci grafik, tabel czy wzorów matematycznych. Skuteczność pracy tego typu oprogramowania ciągle rośnie. Na bazie tej technologii powstają zaawansowane rozwiązania, m.in. takie, jak: aplikacja *Retrievr*, *MindFinder* czy wyszukiwarka w serwisie *Modista.com*.

Segmentacja z kolei bazuje na wyborze reprezentatywnych fragmentów obrazów celem ich porównywania z wcześniej wybranym wzorcem. Chodzi tutaj o stworzenie języka obrazu, definiowanie którego opiera się na słowniku wizualnym (ang. *visual vocabulary*). Słownik wizualny ma za zadanie określać autonomiczne wizualnie figury (fragmenty obrazu), które zostają nazwane *wzorcami wizualnymi*. Elementy reprezentatywne są wybierane w trzech płaszczyznach: horyzontalnej, diagonalnej i wertykalnej. Na podstawie tak utworzonych wzorców, ich powiązań i relacji algorytm jest w stanie rozpoznawać kształty bardziej złożone i w konsekwencji odpowiadać na zapytanie wizualne. Słownik wizualny nie odwołuje się do słowa pisanego, skojarzeń czy wyobrażeń o obrazach.

Wdrożeniami tego typu systemów zainteresowały się w szczególności nauki biomedyczne i techniczne. Powstają komercyjne rozwiązania wśród setek tysięcy obrazów odnajdujące te, w których występują zadane fragmenty. W naukach medycznych powstał na przykład system *ViVo* (Bhattacharya, 2005), który jest w stanie rozpoznawać kształty poszczególnych narządów, a także porównywać ich modele trójwymiarowe ze zdjęciami z interwencji lekarskich. Innym przykładem może być wyszukiwanie zamiennych części samochodowych po wskazaniu zdjęć, na których widnieją jedynie kształty albo też fragmenty kształtów konkretnych podzespołów samochodowych. Z racji przepisów prawnych części te nie mogą mieć zastrzeżonych, oryginalnie fabrycznych oznaczeń producentów.

Ważnym elementem obrazu, który ma odzwierciedlenie w jego reprezentacji słownej, wektorowej lub innej, jest barwa. „W wyniku badań stwierdzono, iż w sytuacji wystarczającego oświetlenia człowiek w pierwszej kolejności reaguje na jasność i kolor, następnie na kształt i ruch, a na końcu na teksturę i inne cechy” (Forczmański, 2005). Deskryptorami graficznymi w wyszukiwarkach wizualnych są dominanty barwne, histogramy i modele (przestrzenie) barw. Są trzy możliwości: wyszukiwanie kolorami, które łączymy z nastrojem, wyszukiwanie kolorem, który występuje w poszukiwanym obrazie oraz identyfikacja barwy z elementem wyszukiwania. W przypadku wyszukiwania kolorem, który występuje w poszukiwanym obrazie, narzędziem do wyszukiwania staje się cyfrowa paleta barw, oparta na technikach procentowego zliczania występowania wybranych kolorów w obrazach. Proces wyszukiwania barw polega na wskazaniu z podstawowej palety barw koloru, ewentualnie kolorów, których szukamy w bazie obrazów. Co więcej, niektóre narzędzia

dopuszczają, aby użytkownik sam określił procentowy udział każdego z kolorów, które chce mieć w wyszukany obrazie. Takie prototypowe rozwiązanie stworzyła na przykład grupa inżynierów z Toronto. Firma TinEye uruchomiła projekt Multicolor Search, który korzysta z obrazów udostępnionych na licencjach Creative Commons z serwisu Flickr. Wyszukiwarka kolorów ma specjalnie zaprojektowany algorytm, współpracujący z większością formatów plików graficznych dostępnych w Internecie. Są to m.in. JPEG (zamiennie można używać JPG), PNG i GIF. Paleta barw RGB umożliwiającą przeszukiwanie bazy obrazów jest automatycznie przeliczana na wartości szesnastkowe (HEX). W procesie indeksacji obrazów nie jest używana żadna instrukcja znakowania kolorów słowami kluczowymi. Narzędziem wyszukiwawczym jest bezpośrednio wybrany przez użytkownika kod koloru. Przeszukiwany obraz posiada specjalne informacje wizualne, charakterystyczne tylko dla niego. Zbiór tych danych twórcy projektu porównują do niepowtarzalnych linii papilarnych na ludzkich dłoniach. Ta niepowtarzalność odnosi się do sposobu rozmieszczenia wybranych barw i ich odcieni na wyszukiwanych obrazach. Kolejnym krokiem w procesie wyszukiwawczym jest tworzenie mapy kolorów. System określa kody poszczególnych kolorów wraz z ich procentowym udziałem w całym obrazie. Na tym etapie użytkownik ma możliwość ingerencji w wyniki końcowe. Można uwzględnić, albo też usunąć, kolor dominujący powiązany z tłem we wczytywanym obrazie. Dzięki takiemu rozwiązaniu w fazie końcowej system bardzo szybko i trafnie odnajduje obrazy podobne kolorystycznie.



Rys. 3. The infinite colors of search. Źródło: Morville & Callender, 2010, 16

Barwa w obrazie łączy się także z jej identyfikacją z elementem wyszukiwania i ma miejsce w wizualizacji marek i produktów (Rys. 3). Określony kolor kojarzy się z zasobem, do którego się odnosi. Naturalnie, proponowana organizacja informacji za pomocą kolorów

nie będzie kompletna z racji kontekstu, w jakim te dane się znajdują. Innym przykładem będzie kojarzenie konkretnie wyszukiwanego koloru z marką, firmą, produktem czy usługą internetową. W przestrzeni sieciowej dochodzi do nieformalnego zastrzegania barw do konkretnej usługi czy marki internetowej. Przykładem może być kojarzenie barwy ciemnoniebieskiej z portalem społecznościowym Facebook, edukacji internetowej rozumianej jako e-learning z kolorem zielonym, handlu elektronicznego z barwą żółtą, a wyszukiwania – z pomarańczową. Atlas barw internetowych jest otwarty, a identyfikacja wizualna usług i branż internetowych rozwija się bardzo dynamicznie.

Kolejnym parametrem opisującym obraz cyfrowy jest jego jasność. Dominanty jasności można zobrazować za pomocą histogramu, czyli graficznego wykresu jasnych, średnich i ciemnych partii obrazu. Rozkład ten jest podstawowym narzędziem służącym do maszynowego uczenia rozpoznawania kształtów. To na podstawie wyodrębnienia różnych jasności i kontrastów komputer jest w stanie wyróżnić i rozdzielić (posegmentować) elementy obrazu. Zasada ta służy do tworzenia algorytmów rozpoznających i wyszukiwujących obrazy, dokonujących autopoprawek zdjęć wykonanych amatorskimi aparatami, umożliwia zaawansowaną korektę (obróbkę) jasności poszczególnych partii obrazu w programie graficznym. Histogram nie charakteryzuje obrazu w sposób jednoznaczny. Wiele różniących się od siebie obrazów może mieć identyczne wartości przebiegu zmienności światła w całym przedstawieniu obrazowym. Nie jest też możliwe odtworzenie obrazu cyfrowego tylko i wyłącznie na podstawie jego histogramu. Technicznie rzecz ujmując, histogramy mogą przybierać zwizualizowane formy dwu- bądź trójwymiarowe. Jeśli nałożymy na siebie trzy histogramy składowych RGB jednego pliku cyfrowego, to umożliwimy komputerowi porównanie obrazów z innymi, które posiadamy w bazie danych. Zasada porównywania trójkanałowych histogramów jest najprostszą metodą wyszukiwania obrazów wizualnie podobnych do wyjściowego. Zaletą histogramów jest ich niska wrażliwość na operacje obrotów, przesunięć i skalowania oraz niska złożoność obliczeniowa. Wyszukiwanie odbywa się według atrybutów formalnych (kolorowe piksele), a nie z wykorzystaniem metadanych czy warstwy zawartości.

#### 4. Wyszukiwanie kontekstowe

Najbardziej zaawansowanym sposobem wyszukiwania materiałów wizualnych jest wyszukiwanie w kontekście. Na podstawie wspomnianego wcześniej słownika wizualnego powstaje nowy typ relacji między obiektami widocznymi na materiałach wizualnych. Wyszukiwanie kontekstowe jest modelem, który sam się rozwija przez wdrażanie coraz to bardziej zaawansowanych zapytań w postaci obrazów. Można zaryzykować stwierdzenie, że wyszukiwanie kontekstowe jest przyszłością wyszukiwania materiałów wizualnych w ogóle. Odpowiada na problem masowego wzrostu popularności obrazów, fotografii, grafik itp. funkcjonujących w sieci, bo nie przekłada języka obrazu na mowę. Na bazie tych rozważań i udoskonalonej wizji Visual Vocabulary powstał nowy sposób opisu treści pod nazwą Visual Memex. W języku tym koduje się zarówno wygląd obiektów, jak i relacje między nimi bez podania źródła, tak jak jest to stosowane w obecnych rozwiązaniach. Algorytm jest w stanie dokładnie wskazać elementy na zdjęciu, ich wzajemne relacje i kontekst przestrzenny, w którym się znajdują. Dzięki temu możliwe jest rozpoznawanie obrazów malarskich, zdjęć

fotograficznych wykonanych w różnej technice, obiektów muzealnych zarówno dwu- jak i trójwymiarowych, grafik, szkiców i rysunków jako jednorodnych obiektów posiadających tę samą treść dostępną w określonym kontekście. Jest to kluczowe podejście w rozumieniu tego, co jest na zdjęciu oraz jaki obiekt przedstawia dany obraz z pominięciem techniki wykonania (sposobu wykonania) czy stopnia transformacji oryginalnego materiału wizualnego. Określenie podobieństwa wizualnego obrazów z różnych dziedzin nie jest możliwe techniką porównywania pikseli. W tym wypadku analizowana jest tzw. zawartość wizualna wyższego poziomu, na przykład w przypadku konkretnej sceny abstrahuje się od pory roku, w jakiej zostało wykonane zdjęcie, natężenia światła, formy – czy jest to obraz malarski, czy też grafika itp. (Shrivrasava et al., 2011).

## 5. Inicjatywy wspomagające rozwój wyszukiwarek wizualnych

Wszystkie wskazane wcześniej techniki komputerowego analizowania i reprezentowania obrazów mimo ich licznych i bezdyskusyjnych zalet zyskują, gdy do pracy z zasobem zostaje włączony człowiek. Im więcej wzorców wizualnych, tym doskonalszy system, stąd inicjatywy angażowania społeczności internautów do wzbogacania bazy wiedzy wyszukiwarek wizualnych. Przykładem takiej praktyki są choćby dwa projekty – Flickr i „Kto jest na Foto”. Innym przykładem jest system How Old<sup>3</sup> (Rys. 4) będący samouczącym się systemem eksperckim określającym wiek i płeć osoby przedstawionej na zdjęciu i wykorzystującym do nauki przykłady wczytywane przez internautów. W polu *Get a human option*, jeśli użytkownik ma odmienne zdanie co do wyniku, który podaje system, można do wczytanego przez siebie zdjęcia dodać poprawny wiek osoby, która na nim widnieje. Jeśli chodzi o identyfikację płci, to praktycznie jest ona rozpoznawana bezbłędnie.

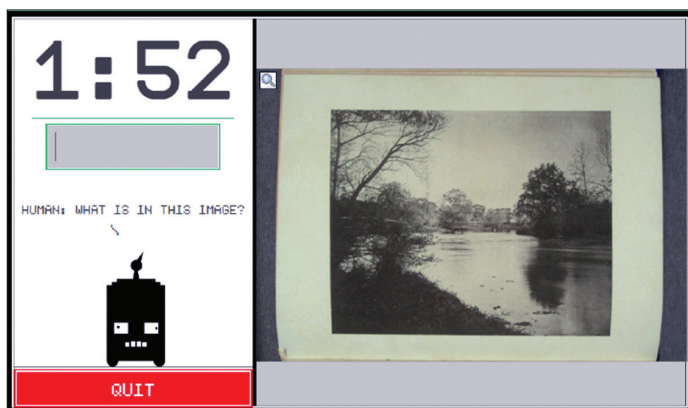


Rys. 4. Szacowane wyniki wieku osoby na wczytanym do systemu portrecie. W tym przypadku jest to słynna *Mona Lisa*. Wczytanie dwóch bardzo podobnych zdjęć skutkuje podaniem zbliżonych, choć nie identycznych wyników. Źródło: <https://how-old.net> [30.01.2019]

<sup>3</sup> Zob. <http://how-old.net>

Inną aktywnością jest uruchamianie „gier metadanych”, na przykład Zen Tag czy Portrait Tag stworzone dla Muzeum Sztuki w Monachium. Aplikacje te są dostępne zarówno w postaci przeglądarkowej, jak i na urządzenia mobilne. Mają one w interaktywny sposób pobudzić kreatywność użytkowników, spowodować chęć rywalizacji w tworzeniu coraz bogatszych treści opisujących cyfrowe obrazy. Internauci najczęściej nadają słowa kluczowe (tagi) opisujące fotografię (kolekcję fotografii), tłumaczą fragmenty zeskanowanego tekstu na inny język lub „nazywają” wyświetlane przez komputer fragmenty obrazów. Projekty przeznaczone dla społeczności internetowej takie jak Open Image Collection czy Library Flickr Commons oferują z kolei możliwości sugerowania użytkownikowi tagów, które już ktoś wcześniej dodał. Mechanizm ten jest także dostępny w aplikacji Artigo<sup>4</sup>, gdzie użytkownik sprawdza sugerowane opisy obrazów. Dodatkowo, aplikacja pyta o relacje treściowe elementów obrazu z określeniami słownikowymi. Użytkownik ma za zadanie odpowiedzieć twierdząco bądź zanegować wybór spośród możliwych odpowiedzi.

Innym przykładem jest aplikacja o nazwie Metadata Games, gdzie to użytkownik informuje algorytm aplikacji, co widzi na ilustracji (Rys. 5). Komputer rozróżnia złożoność fraz wpisywanych przez internautę. Zadawane przez niego pytania początkowo są bardzo proste. Użytkownik, po udzieleniu poprawnych odpowiedzi (pokrywających się z posiadaną bazą odpowiedzi robota), przechodzi do etapu trudniejszego, czyli sugerowania fraz bardziej złożonych, których nie ma jeszcze w bazie danych. Za każdą odpowiedź gracz otrzymuje punkty zgodnie z zasadą: im mniej powtarzalna odpowiedź, tym większa liczba punktów.



Rys. 5. Metadata Games. Źródło: <http://play.metadatagames.org/stupidrobot> [30.01.2019]

## 6. Podsumowanie

Wydaje się, że dziś optymalnym rozwiązaniem problemów wyszukiwania fotografii byłby system hybrydowy, łączący opis tworzony przez profesjonalistę (archiwistę, bibliotekarza, muzealnika) z treściami tworzonymi oddolnie (kolektywnie) przez internautów (tagowanie) oraz z rezultatami komputerowej analizy obrazu. Opis obiektu wizualnego byłby tworzony

<sup>4</sup> Zob. <http://www.muenchen.de/int/en/events/art.html>

równoległe na różnych płaszczyznach, co pozwoliłoby na zmniejszenie ryzyka powstawania błędnych i niezrozumiałych opisów oraz na wzrost możliwości wyszukiwania wizualnego z uwzględnieniem potrzeb użytkowników. Sformułowany przez nas wniosek wprawdzie nie jest odkrywczy, ale zawiera ważną presupozycję, iż podejście tradycyjne, metadane i paradygmat sieciowy nie tworzą alternatywy, a przeciwnie – są komplementarne. Usprawnienie technologii komputerowych zależy przede wszystkim od bogactwa bazy wzorców dostarczanych do rozpoznawania. Dlatego tak istotne jest projektowanie systemów, które w sposób ciągły będą się uczyły nowych figur systematycznie pozyskiwanych z przykładowych obrazów. Współczesna technologia cyfrowa dysponuje dużym potencjałem adaptacyjnym. Kreatywny sposób wykorzystania narzędzi analizujących obrazy cyfrowe może pozwolić na znaczący postęp w dziedzinie wyszukiwania, zwłaszcza że fotografia cyfrowa staje się nie tylko rodzajem ikony, ale też znakiem czasów, wpisując się w kontekst komunikacji wizualnej w wielokanałowej komunikacji międzyludzkiej. Incepcjonizm (rozpoznawanie wyuczonych wzorców obrazowych na innych obrazach) przyspiesza w miarę nabierania przez algorytmy (na przykład Tensor Flow) „doświadczenia” w rozpoznawaniu i uczeniu się wzorców. Eksperymenty z zakresu sztucznej inteligencji prowadzone w przestrzeni wizualnej Internetu dają coraz bardziej interesujące i spektakularne rezultaty. Pozostaje mieć nadzieję, że ten nurt przemian nie ominie bibliotek i ich zasobów fotografii.

## Bibliografia

- Baca, M., ed. (2002). *Introduction to art image access: Issues, tools, standards, strategies* [online], [01.02.2019]. [http://www.getty.edu/research/publications/electronic\\_publications/intro\\_aia/index.html](http://www.getty.edu/research/publications/electronic_publications/intro_aia/index.html)
- Baca, M., Harpring, P., Lanzi, E., McRae, L., Whiteside, A. B. (2006). *Cataloging cultural objects: A guide to describing cultural works and their images* [online], [01.02.2019], [http://cco.vrafoundation.org/index.php/toolkit/cco\\_pdf\\_version/](http://cco.vrafoundation.org/index.php/toolkit/cco_pdf_version/)
- Bal, M. (1991). *Reading "Rembrandt": Beyond the word-image opposition: the Northrop Frye lectures in literary theory*. New York, Cambridge, Amsterdam: University Press.
- Beghtol, C. (1986). Bibliographic classification theory and text linguistics: Aboutness analysis, intertextuality and the cognitive act of classifying documents. *Journal of Documentation*, 42(2), 84–113.
- Bhattacharya, A., Ljosa, V., Jia-Yu Pan, Verardo, M. R., Hyungjeong Yang, Faloutsos, C., Singh, A. K. (2005.). ViVo: Visual Vocabulary Construction for Mining Biomedical Images. In: *Fifth IEEE International Conference on Data Mining (ICDM'05)* (50–57). Houston: IEEE.
- Clough, P. D., Ireson, N., Marlow, J. (2009). Extending domain-specific resources to enable, semantic access to cultural heritage data. *Journal of Digital Information*, 10 (6), [online], [01.02.2019], <https://journals.tdl.org/jodi/index.php/jodi/article/view/698/578>
- Corso, J. (2014). *Semantic knowledge. Generalized image understanding with probabilistic ontologies and dynamic adaptive graph hierarchies* [online], [30.01.2019], <http://www.cse.buffalo.edu/~jcorso/r/career/index.html>
- Domański, M. (2010). *Obraz cyfrowy: Reprezentacja, kompresja, podstawy przetwarzania. Standardy JPEG i MPEG*. Warszawa: Wydaw. Komunikacji i Łączności.
- Forczmański, P., Szeptycki, P. (2005). *Indeksowanie dużych zbiorów obrazów* [online], [30.01.2019], [http://pforczmanski.zut.edu.pl/pub/2005\\_wi.pdf](http://pforczmanski.zut.edu.pl/pub/2005_wi.pdf)
- Geertz, C. (2005). *Interpretacja kultur. Wybrane eseje*. Kraków: Wydaw. UJ.
- Gevers, T., Smeulders, A. W. M. (2003). *Image search engines: An overview*, [online], [30.01.2019], <http://staff.science.uva.nl/~gevers/pub/overview.pdf>

- Jørgensen, C. (2003). *Image retrieval: Theory and research*. Lanham: Scarecrow Press.
- Jørgensen, C., Jaimes, A., Benitez, A. B., Chang, S-F. (2001). A conceptual framework and empirical research for classifying visual descriptors. *Journal of the American Society for Information Science and Technology*, 52(11), 938–947.
- Jørgensen, C. (1999). Access to pictorial material: a review of current research and future prospects. *Computer and the Humanities*, 33(4), 293–318.
- Jørgensen, C. (1998). Attributes of images in describing tasks. *Information Processing & Management*, 34(2/3), 161–174.
- Miller, M., Wornbard, M. (2009). Fotografie w zbiorach cyfrowych – problemy z opracowaniem formalnym i rzeczowym na przykładzie Biblioteki Cyfrowej Politechniki Warszawskiej. *Przegląd Biblioteczny*, 77(2), 201–218.
- Morville, P., Callender, J. (2010). *Search Patterns: Design for Discovery*. Sebastopol: O'Reilly Media.
- Rygiel, P. (2012). Teoria i praktyka opracowania rzeczowego dokumentów ikonograficznych w dużych kolekcjach bibliotecznych. *Zagadnienia Informatyki Naukowej*, 98(1), 77–94.
- Shatford-Layne, S. (1994). Some issues in the indexing of images. *Journal of the American Society for Information Science*, 45(8), 583–588.
- Shatford, S. (1986). Analyzing the subject of a picture: A theoretical approach. *Cataloging & Classification Quarterly* 6(3), 39–62.
- Staszczak, R. Znajdź mnie, jeśli potrafisz – czyli jak znaleźć obrazy w Sieci [online], [30.01.2019], <http://www.chip.pl/artykuly/porady/2010/06/znajdz-mnie-jesli-potrafisz#ixzz47pHZv6c5>
- Svenonius, E. (1994). Access to non-book materials: the limits of subject indexing for visual and aural languages. *Journal of the American Society of Information Science*, 45(8), 600–606.
- Tabakowska, E. (2006). *Ikoniczność znaku: słowo – przedmiot – obraz – gest*. Kraków: Universitas.
- Talbierska, J. (2009). Zbiory artystyczne, zbiory specjalne, czy zbiory ikonograficzne? Typologia i funkcje. *Annales Universitatis Wratislaviensis*, 3167, *Bibliotekoznawstwo*, 28, 19–38.
- Wilson, T. D. (2000). Human information behaviour. *Informing Science. The International Journal of an Emerging Transdiscipline*, 3(2), 49–56.
- Włodarczyk, B., Woźniak-Kasperek, J. (2017). *Język haseł przedmiotowych Biblioteki Narodowej: od analizy dokumentu do opisu przedmiotowego*, Warszawa: Wydaw. SBP.

---

## Photography in the Library and on the Web: Two Perspectives on Representation and Retrieval

### Abstract

**Purpose/Thesis:** The purpose of the article is to present two approaches to photographs representation and retrieval: the traditional one, library and metadata-based and the automatic one, known as the network approach. The discussion is focused on indicating the possibility of the mutual complementation of both approaches in order to create a new information and search/retrieval quality in multi-channelled human communication.

**Approach/Methods:** The method of literature analysis and critical literature review was used.

**Results and conclusions:** Today the optimum solution for the problems of photographs representation and retrieval would be a hybrid system combining descriptions created by professionals with the content generated by Internet users (tagging) and the results of computer image analysis. The description of a visual object created at different levels would reduce the risk of producing erroneous, incomprehensible and incomplete descriptions while enhancing the possibility of visual objects retrieval, with the individual needs of users taken into consideration.

**Originality/Value:** The article discusses the issues of computer image analysis, in particular photographs analysis, and its usefulness for information and search purposes: these issues are still underrepresented in Polish book and information science literature.

**Keywords**

Computer image analysis. Image representation. Library. Photographs. Retrieval. Searching.

---

*KAMIL STĘPIEŃ, doktor nauk humanistycznych w dyscyplinie bibliologia i informatologia, adiunkt w Zakładzie Informatologii Instytutu Informacji Naukowej i Bibliotekoznawstwa UMCS. Zainteresowania badawcze: wizualizacja informacji, wyszukiwanie wizualne, funkcjonowanie fotografii w przestrzeni sieciowej i bibliotecznej, tagowanie i folksonomia. Ważniejsze publikacje: Folksonomie, czyli społecznościowe opisywanie treści (Warszawa, Wydaw. SBP, 2010); Wyszukiwanie obrazów i wyszukiwanie obrazem w przestrzeni internetowej. [W:] Z. Osiński i in. (red.), Biblioteka, Książka, Informacja i Internet (47–59) (Lublin: Wydaw. UMCS, 2015); Komunikacja fotograficzna. Interpretacja fotografii, Kultura i Historia 2018, nr 34, 162–186.*

*Kontakt z autorem:*

*kamil.stepien@umcs.pl*

*Zakład Informatologii, Instytut Informacji Naukowej i Bibliotekoznawstwa*

*Uniwersytet Marii Curie-Skłodowskiej w Lublinie*

*Plac Marii Skłodowskiej-Curie 4*

*20-031 Lublin*

*JADWIGA WOŹNIAK-KASPEREK, profesor nauk humanistycznych, kierownik Katedry Bibliografii i Dokumentacji na Wydziale Dziennikarstwa, Informacji i Bibliologii UW. Zainteresowania badawcze: informacja terminologiczna; teoria bibliografii; reprezentacja, organizacja i wyszukiwanie informacji. Ważniejsze publikacje: Problemy współpracy narzędzi wyszukiwawczych bibliografii specjalnych z uniwersalnymi systemami organizacji wiedzy, Śląski Kwartalnik Historyczny Sobótka 2016, 71(3), 39–44; z Bartłomiejem Włodarczykiem, Język haseł przedmiotowych Biblioteki Narodowej: od analizy dokumentu do opisu przedmiotowego (Warszawa: Wydaw. SBP, 2017); Kłamstwo i metoda: na przykładach reprezentacji treści w katalogach bibliotecznych i w bibliografiach (Warszawa: BEL Studio, 2018).*

*Kontakt z autorką:*

*jbwozniak@uw.edu.pl*

*Katedra Bibliografii i Dokumentacji, Wydział Dziennikarstwa, Informacji i Bibliologii*

*Uniwersytet Warszawski*

*ul. Nowy Świat 69*

*00-046 Warszawa*