**Przemysłowy system wizyjny do detekcji drobnych przedmiotów złotniczych i jubilerskich**

**Grupa laboratoryjna 3**

Stanisław Horna **241202**

Przemysław Kowalski **241214**

Spis treści

[**1.** **Temat projektu** 2](#_Toc166704953)

[**2.** **Wstęp** 3](#_Toc166704954)

[**3.** **Materiały i metody** 4](#_Toc166704955)

[3.1. Materiały badawcze 4](#_Toc166704956)

[3.2. Opis stanowiska symulującego taśmociąg 5](#_Toc166704957)

[**3.2.1. Taśmociąg** 5](#_Toc166704958)

[**3.2.2. Kamera** 6](#_Toc166704959)

[**3.2.3. Oświetlenie** 7](#_Toc166704960)

[**3.2.4. Schemat stanowiska** 8](#_Toc166704961)

[3.3. Metody 9](#_Toc166704962)

[**3.3.1. Wprowadzanie obrazu** 9](#_Toc166704963)

[**3.3.2. Transformacja obrazu** 9](#_Toc166704964)

[**3.3.3. Detekcja obiektów** 18](#_Toc166704965)

[**3.3.4. Zliczanie obiektów** 19](#_Toc166704966)

[**4.** **Wyniki** 20](#_Toc166704967)

[**5.** **Wnioski** 20](#_Toc166704968)

[**6. Literatura** 20](#_Toc166704969)

# **Temat projektu**

Przemysłowy system wizyjny, który w czasie rzeczywistym, wykona detekcję, a następnie klasyfikację wyrobów jubilerskich umieszczonych na przemieszczającym taśmociągu.

# **Wstęp**

Rozpoznawanie biżuterii jest złożonym zadaniem ze względu na różnych stylów i wzorów akcesoriów. Dokładne opisy różnych akcesoriów jest czymś, co w dzisiejszych czasach może być osiągnąć tylko eksperci w dziedzinie biżuterii. W niniejszej pracy podejście do rozpoznawania biżuterii przy użyciu technik techniki wizyjne i podpisywanie obrazów, próbując symulować to eksperckie ludzkie zachowanie analizowania akcesoriów. Proponowana metodologia polega na wykorzystaniu różnych modeli podpisywania obrazów do wykrywania klejnotów z obrazu i generowania naturalnego opisu opis akcesorium w języku naturalnym. Następnie opis ten jest jest również wykorzystywany do klasyfikowania akcesoriów na różnych poziomach szczegółowości. Wygenerowany podpis zawiera szczegóły, takie jak rodzaj klejnotu, kolor, materiał i projekt. Aby zademonstrować skuteczność proponowanej metody w dokładnym rozpoznawaniu różnych typów klejnotów, zbiór danych składający się z obrazów akcesoriów należących do sklepów jubilerskich w C´ordobie (Hiszpania). Po testowaniu różnych zaprojektowanych architektur podpisywania obrazów, metoda w ostatecznym modelu osiąga dokładność napisów na poziomie 95%. Proponowana metodologia ma potencjał do wykorzystania w różnych aplikacjach takich jak handel elektroniczny biżuterią, zarządzanie zapasami lub automatyczne rozpoznawanie automatyczne klejnotów w celu analizy gustów i statusu społecznego ludzi [[1]](J.%20M.%20Alcalde-Llergo,%20E.%20Yeguas-Bolívar,%20A.%20Zingoni%20and%20A.%20Fuerte-Jurado,%20%22Jewelry%20Recognition%20via%20Encoder-Decoder%20Models,%22%202023%20IEEE%20International%20Conference%20on%20Metrology%20for%20eXtended%20Reality,%20Artificial%20Intelligence%20and%20Neural%20Engineering%20(MetroXRAINE),%20Milano,%20Italy,%202023,%20pp.%20116-121,%20doi:%2010.1109/MetroXRAINE58569.2023.10405609.%20keywords:%20%7bImage%20recognition;Natural%20languages;Neural%20engineering;Metrology;Behavioral%20sciences;Task%20analysis;Testing;Image%20Captioning;Classification;Object%20Detection;Jewelry;Deep%20Learning;Human%20Behavior%7d,).

Nasz projekt różni się od rozwiązania znalezionego w Internecie m.in.:

* możliwością detekcji wielu różnych obiektów z różnych kategorii na jednym rysunku,
* możliwością podglądu statystyk tj. ile poprawnie zliczono obiektów z danej kategorii.

Do realizacji projektu zostało przygotowane m.in. stanowisko (taśmociąg) oraz wykorzystano wiele różnych metod przetwarzania obrazu jak:

* progowanie,
* rozmycie metodą Gaussa,
* detekcja krawędzi metodą Canny,
* morfologiczne domknięcie krawędzi,
* detekcja obiektów.

# **Materiały i metody**

## 3.1. Materiały badawcze

Przedmioty wrzucane na taśmę to drobne przedmioty złotnicze i jubilerskie w kategoriach:

* naszyjniki (złote, srebrne):  
  

Rys. 1 Złoty i srebrny naszyjnik

* pierścionki (obrączki i zdobieniami np. kamieniem szlachetnym):  
  

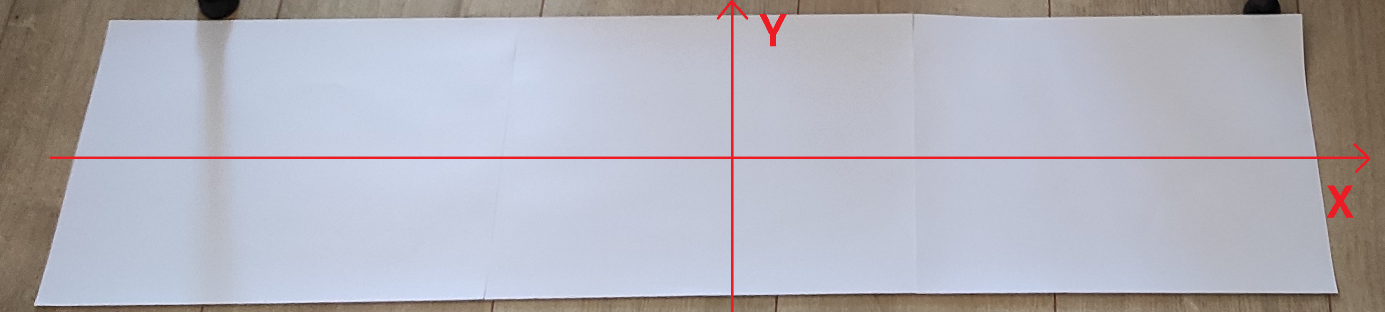
Rys. 2 Obrączki i pierścień z kamieniem szlachetnym

* kolczyki:  
  

Rys. 3 Kolczyki zwykłe i z zdobieniami np. kamieniem szlachetnym

## 3.2. Opis stanowiska symulującego taśmociąg

### **3.2.1. Taśmociąg**



Rys. 4 Taśmociąg

Drobne przedmioty złotnicze i jubilerskie będą wrzucane na taśmociąg, który będzie poruszał się ze stałą prędkością. Pas transmisyjny taśmociągu powinien być gładki wykonany z gumy lub papieru w kolorze białym (***#ffffff*** w systemie heksadecymalnym).

Taśmociąg będzie przesuwany w osi ***X*** (poziomej)z prędkością .

Dzięki temu uzyskujemy lepszą reprezentację kolorów oraz kontrast pomiędzy obiektami a tłem.

Przedmioty wpadają na taśmociąg w sposób, który umożliwia ich przyleganie do siebie, ale nie nakładanie się jednego przedmiotu na drugi, czy umieszczanie jednego w drugim.

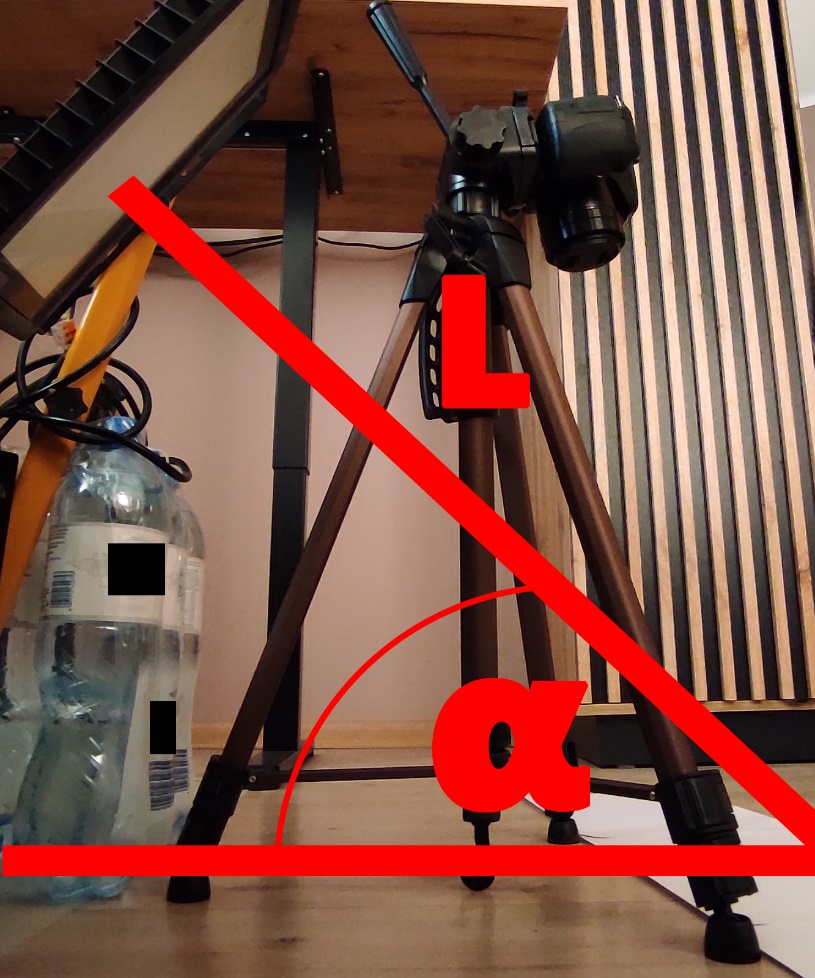
### **3.2.2. Kamera**



Rys. 5 Ustawienie kamery

Zainstalowana kamera będzie rejestrować przedmioty transportowane przez taśmociąg pod kątem **α=85°** z odległości **L=40cm**, która nagrywa minimum w rozdzielczości **1920x1080** (FullHD) oraz **30** klatkach na sekundę (30fps). Obraz powinien zostać odpowiednio pomniejszony przed rozpoczęciem analizy poszczególnych klatek.

### **3.2.3. Oświetlenie**

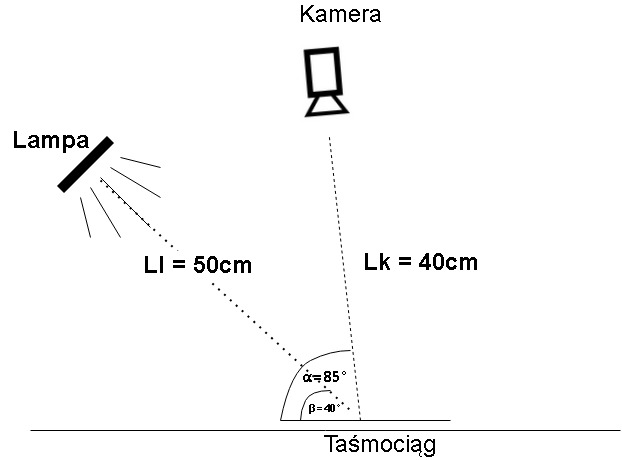


Rys. 6 Ustawienie lampy

Taśmociąg będzie oświetlony jedną lampą generującą **8550-9000lm**. Lampa będzie ustawiona w odległości **L=50cm** od taśmy przesuwającej przedmioty pod kątem **α=40°**.

### **3.2.4. Schemat stanowiska**

Stanowisko można opisać według poniższego schematu:



Rys. 7 Schemat techniczny stanowiska

gdzie:

* **Lk** – odległość kamery od taśmociągu,
* **α** – kąt ustawienia kamery względem taśmociągu,
* **β** - kąt ustawienia lampy względem taśmociągu,,
* **Ll** - odległość lampy (źródła światła) od taśmociągu,

## 3.3. Metody

Obraz zawierający linia, biały

Opis wygenerowany automatycznie

Rys. 8 Schemat ogólny

### **3.3.1. Wprowadzanie obrazu**

Rozdzielczość przetwarzanego obrazu do rozmiaru 1280x720 pixeli, który pozwala uzyskać większą wydajność programu, oraz w zupełności wystarcza do poprawnej identyfikacji obiektów.

### **3.3.2. Transformacja obrazu**

Została podzielona na poszczególne etapy:

Obraz zawierający tekst, Czcionka, diagram, linia

Opis wygenerowany automatycznie

Rys. 9 Schemat szczegółowy

#### **3.3.2.1. Konwersja na odcienie szarości**

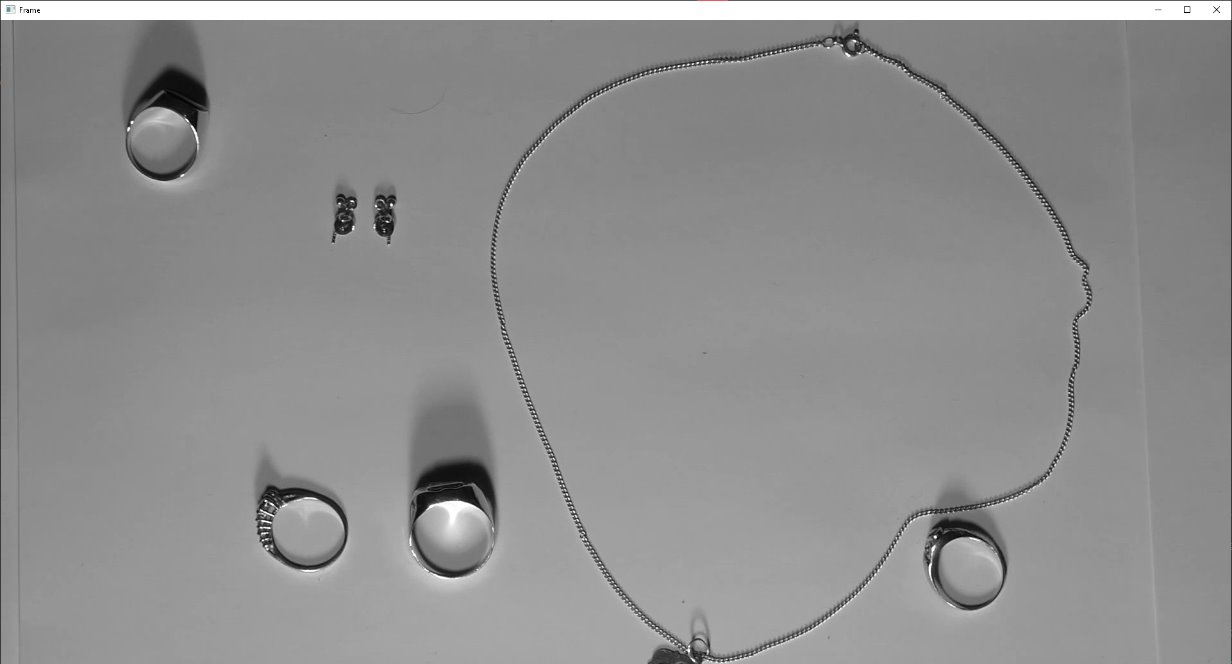
Obraz zawierający tekst, Czcionka, diagram, linia

Opis wygenerowany automatycznie

Rys. 10 Etap 1 - zamiana na odcienie szarości

W pierwszym kroku zamieniamy kolorowy obraz na odcienie szarości. W wyniku tej operacji uzyskujemy:

* obraz jest w skali czerń - biel (0 - 255),
* redukcję szumów,
* operacje na macierzach o dwóch wymiarach, zamiast trzech,
* lepszą detekcję krawędzi i krańców elementów na kolejnych etapach transformacji.



Rys. 11 Efekt po zamianie na odcienie szarości

#### **3.3.2.2. Rozmycie metodą Gaussa**

Obraz zawierający tekst, Czcionka, diagram, linia

Opis wygenerowany automatycznie

Rys. 12 Etap 2 - rozmycie metodą Gaussa

Rozmycie Gaussa to rodzaj filtra rozmywającego obraz, który używa funkcji Gaussa, do obliczenia transformacji dla każdego piksela na obrazie. Formuła funkcji Gaussa w jednym wymiarze to:

W dwóch wymiarach jest to iloczyn dwóch takich funkcji Gaussa, po jednej w każdym wymiarze:

gdzie:

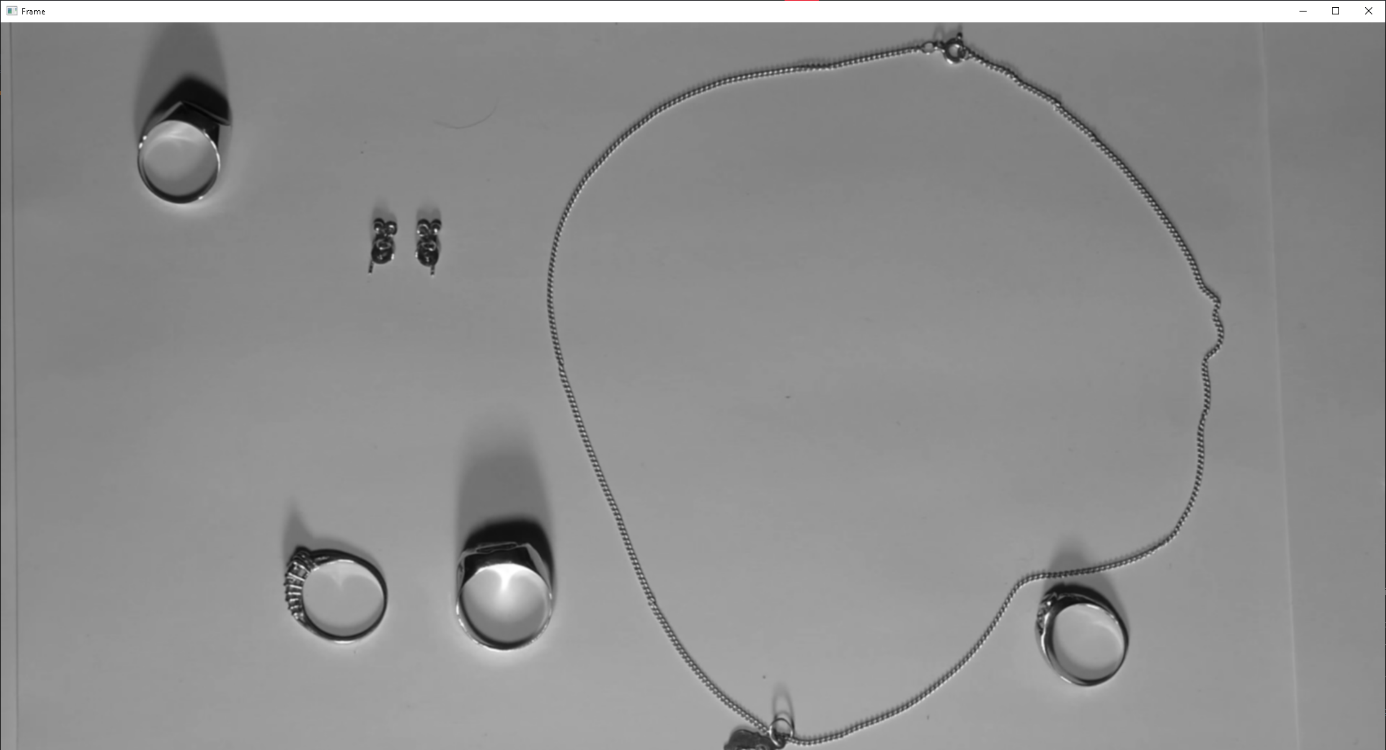
* **x** - odległość od początku na osi poziomej,
* **y** - odległość od początku na osi pionowej,
* **σ** - odchylenie standardowe rozkładu Gaussa.

Ważne jest, aby zauważyć, że początek na tych osiach znajduje się w centrum (0, 0). Po zastosowaniu w dwóch wymiarach, ta formuła produkuje powierzchnię, której kontury są okręgami o rozkładzie Gaussa od punktu centralnego.

Wartości z tego rozkładu są używane do zbudowania macierzy splotu, która jest stosowana do oryginalnego obrazu. Nowa wartość każdego piksela jest ustawiana jako średnia ważona sąsiedztwa tego piksela. Wartość oryginalnego piksela otrzymuje największą wagę (ma najwyższą wartość gaussowską), a sąsiednie piksele otrzymują mniejsze wagi wraz ze wzrostem ich odległości od oryginalnego piksela. Skutkuje to rozmyciem, które zachowuje granice i krawędzie lepiej niż inne, bardziej jednolite filtry rozmywające; zobacz także implementację przestrzeni skali. [[2]](https://en.wikipedia.org/wiki/Gaussian_blur)

W naszym przypadku zastosowaliśmy funkcję Gaussa w dwóch wymiarach podstawiając:

* **x = 1**,
* **y** = **1**,
* **σ = 5**.

****

Rys. 13 Efekt po rozmyciu metodą Gaussa

##### **3.3.2.3. Detekcja krawędzi algorytmem Canny**

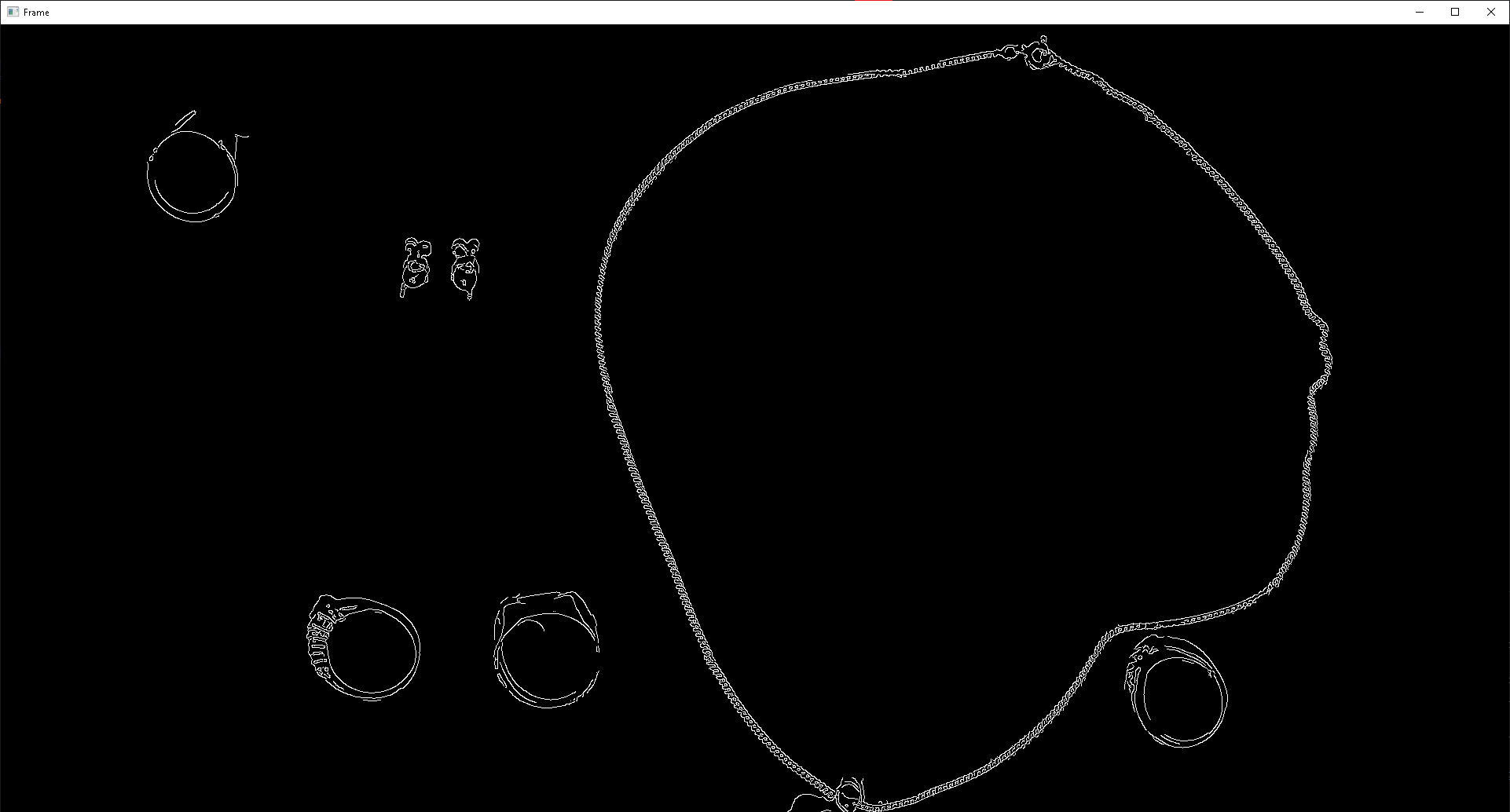
Obraz zawierający tekst, Czcionka, diagram, linia

Opis wygenerowany automatycznie

Rys. 14 Etap 3 - detekcja krawędzie metodą Canny

Detekcja krawędzi metodą Canny to technika, która pozwala na wyodrębnienie użytecznych informacji strukturalnych z różnych obiektów widzenia i znaczne zmniejszenie ilości danych do przetworzenia. Jest to proces wieloetapowy, który składa się z następujących kroków:

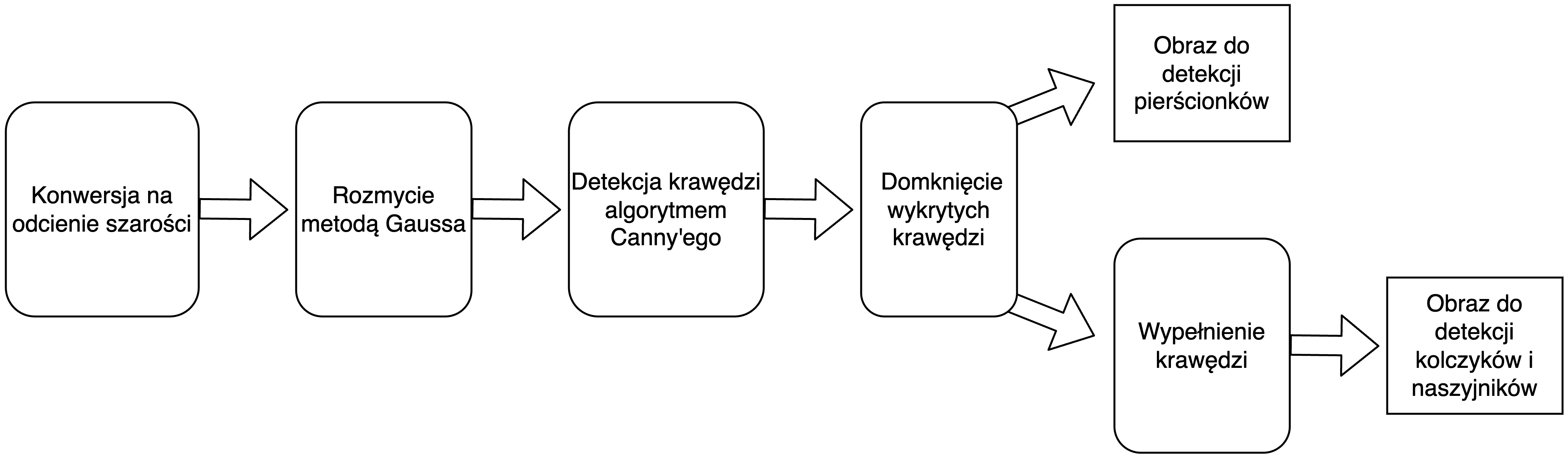
1. **Redukcja szumów**: Ponieważ operacje matematyczne w tej metodzie bazują głównie na pochodnych, wyniki detekcji krawędzi są bardzo wrażliwe na szumy obrazu. Aby pozbyć się szumów, stosuje się rozmycie Gaussa, którą wykonaliśmy w poprzednim etapie.
2. **Obliczanie gradientu**: Krok ten polega na wykrywaniu intensywności krawędzi i kierunku, obliczając gradient obrazu za pomocą operatorów detekcji krawędzi.  
   gdzie **\*** oznacza operację splotu. Na podstawie obliczonych ***Gx*** i ***Gy*** oblicza się kierunek (**G**) i wielkość () gradientu:
3. **Tłumienie maksimum**: W tym kroku, wszystkie piksele, które nie są na maksymalnym gradiencie (nie są najjaśniejsze piksele) są tłumione (zerowane).
4. **Podwójny próg**: Ten krok polega na podziale pikseli na trzy kategorie: piksele silne, piksele słabe i piksele nie krawędziowe.
5. **Śledzenie krawędzi przez histerezę**: W tym etapie, piksele słabe są albo odrzucane, albo zmieniane na piksele silne (krawędzie), w zależności od ich połączenia z innymi pikselami [[3]](https://hbyacademic.medium.com/canny-edge-detection-54bd845940c3)



Rys. 15 Efekt po detekcji krawędzi algorytmem Canny

W rezultacie otrzymujemy obraz w postaci binarnej (gdzie wartość 0 to czerń, a 1 to biel).

##### **3.3.2.4 Domknięcie wykrytych krawędzi**



Morfologiczne domknięcie krawędzi, znane również jako operacja zamknięcia w przetwarzaniu obrazów, jest operacją, która składa się z dwóch etapów:

1. **Dylatacja**: Dylatacja jest pierwszym etapem operacji zamknięcia. Dylatacja obrazu **A** przez element strukturalny **B** jest zdefiniowana jako:

gdzie oznacza przesunięcie elementu strukturalnego **B** do lokalizacji **z**.

1. **Erozja**: Erozja jest drugim etapem operacji zamknięcia. Erozja obrazu A przez element strukturalny B jest zdefiniowana jako:

gdzie oznacza przesunięcie elementu strukturalny **B** do lokalizacji **z**.

Operacja zamknięcia jest zatem zdefiniowana jako dylatacja obrazu, a następnie erozja tego samego obrazu, co można zapisać jako:

Operacja ta jest używana do wygładzania konturów, łączenia wąskich przerw i długich, cienkich zatok, eliminowania małych dziur i wypełniania luk w konturze. [[4]](https://www.cse.iitd.ac.in/~pkalra/csl783/Morphological.pdf)

Ad 1. Przykład dla macierzy **A** o wymiarach 11x11 i macierzy **B** o wymiarach 3x3:  
 0 0 0 0 0 0 0 0 0 0 0

0 1 1 1 1 0 0 1 1 1 0

0 1 1 1 1 0 0 1 1 1 0

0 1 1 1 1 1 1 1 1 1 0

0 1 1 1 1 1 1 1 1 1 0 1 1 1

0 1 1 0 0 0 1 1 1 1 0 1 1 1

0 1 1 0 0 0 1 1 1 1 0 1 1 1

0 1 1 0 0 0 1 1 1 1 0

0 1 1 1 1 1 1 1 0 0 0

0 1 1 1 1 1 1 1 0 0 0

0 0 0 0 0 0 0 0 0 0 0

Dla każdego piksela w macierzy **A**, który ma wartość **1**, nakładana jest macierz **B**, w taki sposób, że środek macierzy **B** jest wyrównany z odpowiednikiem piksela w macierzy **A**.

Rezultatem dylatacji macierzy **A** z macierzą **B** jest macierz o wymiarach 11x11:

1 1 1 1 1 1 1 1 1 1 1

1 1 1 1 1 1 1 1 1 1 1

1 1 1 1 1 1 1 1 1 1 1

1 1 1 1 1 1 1 1 1 1 1

1 1 1 1 1 1 1 1 1 1 1

1 1 1 1 1 1 1 1 1 1 1

1 1 1 1 0 1 1 1 1 1 1

1 1 1 1 1 1 1 1 1 1 1

1 1 1 1 1 1 1 1 1 1 1

1 1 1 1 1 1 1 1 1 0 0

1 1 1 1 1 1 1 1 1 0 0

Ad 2. Przykład dla macierzy **A** o wymiarach 13x13 i macierzy **B** o wymiarach 3x3:

1 1 1 1 1 1 1 1 1 1 1 1 1

1 1 1 1 1 1 0 1 1 1 1 1 1

1 1 1 1 1 1 1 1 1 1 1 1 1

1 1 1 1 1 1 1 1 1 1 1 1 1

1 1 1 1 1 1 1 1 1 1 1 1 1

1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

1 1 1 1 1 1 1 1 1 1 1 1 1

1 1 1 1 1 1 1 1 1 1 1 1 1

1 1 1 1 1 1 1 1 1 1 1 1 1

1 1 1 1 1 1 1 1 1 1 1 1 1

1 1 1 1 1 1 1 1 1 1 1 1 1

Zakładając, że punkt początkowy **B** znajduje się w jego środku, dla każdego piksela w **A** nałóż punkt początkowy **B**, jeśli **B** jest całkowicie zawarty w **A**, piksel zostanie zachowany, w przeciwnym razie zostanie usunięty (wyzerowany).

Rezultatem erozji macierzy **A** z macierzą **B** jest macierz o wymiarach 13x13:

0 0 0 0 0 0 0 0 0 0 0 0 0

0 1 1 1 1 0 0 0 1 1 1 1 0

0 1 1 1 1 0 0 0 1 1 1 1 0

0 1 1 1 1 1 1 1 1 1 1 1 0

0 1 1 1 1 1 1 1 1 1 1 1 0

0 1 1 1 1 1 1 1 1 1 1 1 0

0 1 1 1 1 1 1 1 1 1 1 1 0

0 1 1 1 1 1 1 1 1 1 1 1 0

0 1 1 1 1 1 1 1 1 1 1 1 0

0 1 1 1 1 1 1 1 1 1 1 1 0

0 1 1 1 1 1 1 1 1 1 1 1 0

0 1 1 1 1 1 1 1 1 1 1 1 0

0 0 0 0 0 0 0 0 0 0 0 0 0

W praktyce, dla każdego piksela w obrazie, operacja dylatacji zastępuje piksel maksymalną wartością wszystkich pikseli w sąsiedztwie zdefiniowanym przez element strukturalny. Następnie operacja erozji zastępuje piksel minimalną wartością wszystkich pikseli w sąsiedztwie zdefiniowanym przez ten sam element strukturalny. W efekcie operacja zamknięcia wygładza kontury, łączy wąskie przerwy i długie, cienkie zatoki, eliminuje małe dziury i wypełnia luki w konturze.

Tak przetworzony obraz jest gotowy do rozpoznawanie przedmiotów na taśmie transmisyjnej które powinny być zakwalifikowane do kategorii „pierścionki”.

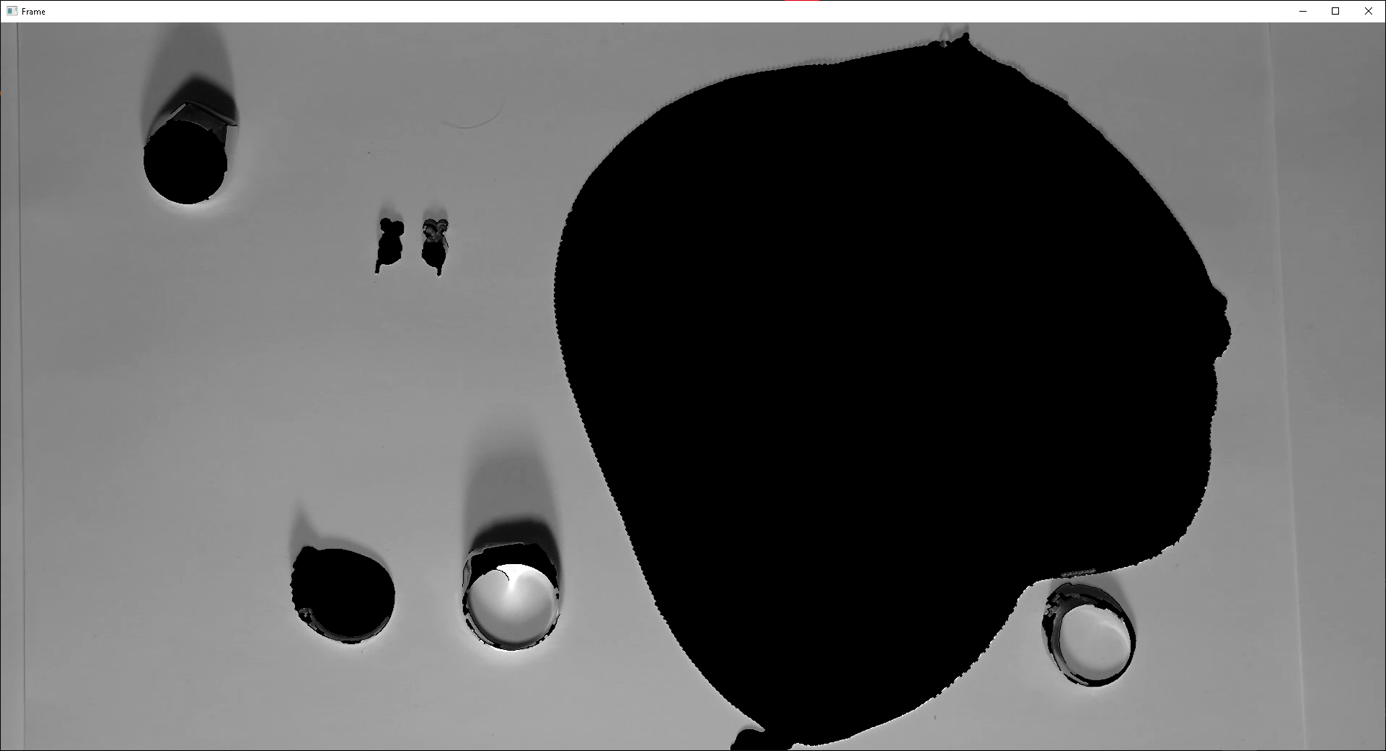
###### **3.3.2.5. Wypełnienie krawędzi**

Obraz zawierający tekst, Czcionka, diagram, linia

Opis wygenerowany automatycznie

**W tym przypadku wykonujemy poszczególne kroki:**

1. **Znajdowanie konturów**: Po operacji morfologicznego zamknięcia krawędzi, używamy metody do odnalezienia konturów na obrazie. Matematycznie, ta operacja polega na znalezieniu ciągłych krzywych na obrazie, które łączą wszystkie punkty o tej samej intensywności. Metoda ta powinna zwraca wszystkie kontury bez uwzględniania ich hierarchii.
2. **Kompresja konturów**: Matematycznie, ta operacja polega na redukcji liczby punktów w konturze, zastępując ciąg punktów linią prostą między pierwszym a ostatnim punktem.
3. **Wypełnianie konturów**: Wszystkie odnalezione kontury wypełniamy kolorem czarnym. Matematycznie, ta operacja polega na wypełnieniu obszarów ograniczonych konturami określonym kolorem.

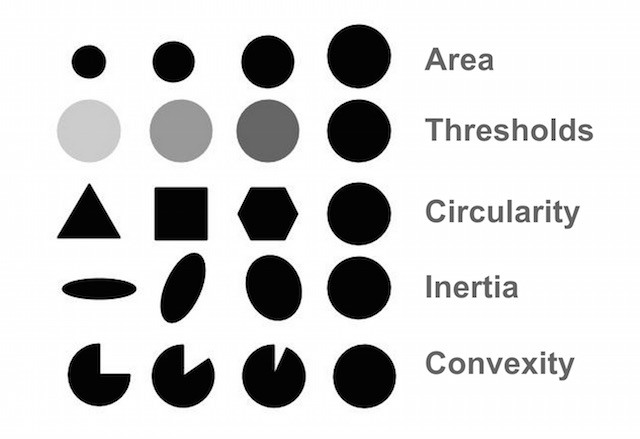


Rys. 18 Efekt ponownego znalezienia krawędzie i ich wypełnienia

### **3.3.3. Detekcja obiektów**

Wszystkie obiekty są wykrywane poprzez *Blob* (Binary Large Object) czyli grupy połączonych pikseli na obrazie, które mają podobne właściwości np. kolor. Detekcja jest wykonywana na podstawie parametrów:

* rozmiaru zajmowanego obszaru (area),
* wypukłości (convexity),
* inercji (inertia),
* kolistości (circularity),
* koloru (color).



Rys. 19 Graficzna reprezentacja parametrów do detekcji obiektów

Wyjaśnienia do poszczególnych parametrów:

* **Area** – obszar zajmowany przez obiekt,
* **Thresholds** – obiekty o poszczególnym odcieniu,
* **Circularity** - kolistość – czyli jak bardzo obiekt jest zbliżony do kształtu okręgu. Im wartość niższa tym obiekt jest bardziej zbliżony do kształtu trójkąta. Parametr opisany jest wzorem:

gdzie:

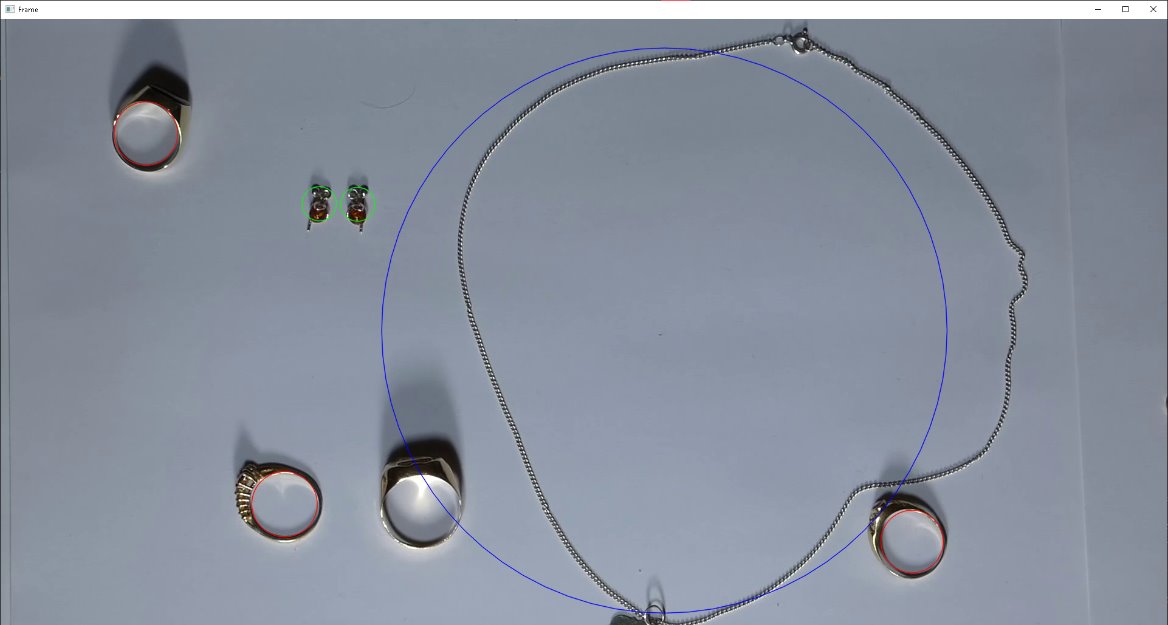
* + **P** – pole powierzchni,
  + **L** – obwód,
* **Convexity** – wypukłość - jest definiowana jako (pole powierzchni bryły / pole jej wypukłego kadłuba). Wypukły kadłub kształtu jest najściślejszym wypukłym kształtem, który całkowicie otacza kształt,
* **Inertia** (inercja) - mierzy, jak wydłużony jest kształt np. dla okręgu.

Dla wszystkich obiektów parametr *kolor* powinien być ustawiony na wartość 0 (czerń), gdyż po wstępnej transformacji obraz jest w odcieniach szarości.

Naszyjniki zajmują najwięcej obszaru oraz mają kształt kulisty.

Pierścionki mają również kształt kulisty oraz zajmują mniej obszaru niż naszyjniki. Detekcja pierścionków jest wykonywana na podstawie okręgu wewnętrznego okręgu (miejsca na palec osoby, która go nosi).

Kolczyki mają różne kształty oraz zajmują jeszcze mniej obszaru niż pierścionki. Detekcja kolczyków wykonywana jest na podstawie kształtu oraz inercji. Dla kolczyków pomijamy detekcję poprzez parametr *wypukłości*.



Rys. 20 Efekt detekcji obiektów

### **3.3.4. Zliczanie obiektów**

TODO (in progress)

# **Wyniki**

Dla obiektów typu pierścionki wyniki były najlepsze. Skuteczność detekcji wynosi ok.90%. Dla kolczyków jest to ok. 80% natomiast dla naszyjników ok. 60%.

# **Wnioski**

Drobne przedmioty złotnicze i jubilerskie odbijają światło co początkowo utrudniało prawidłową detekcję obiektów. Dzięki metodom takimi jak progowanie, rozmycie obrazu metodą *Gaussa*, detekcje krawędzi metodą *Canny* oraz morfologiczne domknięcie krawędzi uzyskaliśmy obraz, który miał zniwelowane efekty odbicia światła. Na bazie zmodyfikowanego obrazu mogliśmy wykonać właściwą detekcję obiektów. Rezultaty naszej metody detekcji obiektów są bardziej skuteczne przy lepszym oświetleniu stanowiska. Po wielu próbach ustawienia pozycji oświetlenia najlepsze rezultaty otrzymywaliśmy gdy lampa była ustawiona w odległości **50cm** od taśmociągu pod kątem **40°**

# **6. Literatura**

[[1] J. M. Alcalde-Llergo, E. Yeguas-Bolívar, A. Zingoni and A. Fuerte-Jurado, "Jewelry Recognition via Encoder-Decoder Models," 2023 IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE), Milano, Italy, 2023, pp. 116-121, doi: 10.1109/MetroXRAINE58569.2023.10405609. keywords: {Image recognition;Natural languages;Neural engineering;Metrology;Behavioral sciences;Task analysis;Testing;Image Captioning;Classification;Object Detection;Jewelry;Deep Learning;Human Behavior},](https://ieeexplore.ieee.org/document/10405609)

[[2] Gaussian blur](https://en.wikipedia.org/wiki/Gaussian_blur)

[[3] Canny edge detection](https://hbyacademic.medium.com/canny-edge-detection-54bd845940c3)

[[4] Morphological.pdf (iitd.ac.in)](https://www.cse.iitd.ac.in/~pkalra/csl783/Morphological.pdf)