

**Akademia Górniczo-Hutnicza  
im. Stanisława Staszica w Krakowie**

---

Wydział Elektrotechniki, Automatyki, Informatyki i Inżynierii Biomedycznej



**Praca magisterska**

Damian Sidor

***Wykorzystanie cienia do odtworzenia pozycji obiektu  
w przestrzeni***

Promotor  
dr inż. Paweł Rotter

Kraków 2013

### **Oświadczenie pracy autora**

Oświadczam, świadomy odpowiedzialności karnej za poświadczenie nieprawdy, że niniejszą pracę dyplomową wykonałem osobiście i samodzielnie i że nie korzystałem ze źródeł innych niż wymienione w pracy.

.....  
Podpis

**AGH**  
**University of Science and Technology in Krakow**

---

Faculty of Electrical Engineering, Automatics, Computer Science and Biomedical  
Engineering



**Master of Science Thesis**

Damian Sidor

*Calculation of the object position in 3D space  
from the shadow*

Supervisor  
Paweł Rotter Ph.D

Kraków 2013

## Spis treści

Wstęp.....	8
Cel i zakres pracy.....	9
Struktura pracy.....	9
1 Tradycyjna stereowizja.....	11
1.1 Geometria epipolarna.....	11
1.2 System stereowizyjny oparty na transformacji płaskiej.....	13
1.3 Systemy stereowizyjne z jedną kamerą.....	14
2 Cień w systemach wizyjnych.....	18
3 Geometria systemu stereowizyjnego wykorzystującego cień obiektu.....	21
3.1 Kalibracja płaszczyzn - transformacja „płaska”.....	21
3.2 Kalibracja światła – wyznaczenie współrzędnych punktu światła lub jego kąta padania.....	24
3.2.1 Źródło światła o promieniach rozchodzących się promieniście – żarówka.....	24
3.2.2 Źródło światła o promieniach równoległych – słońce.....	27
3.3 Wyznaczanie położenia obiektu.....	28
4 Analiza i przetwarzanie obrazu.....	32
4.1 Algorytm przetwarzania obrazu.....	32
4.2 Opis poszczególnych kroków przetwarzania.....	34
4.3 Inne koncepcje przetwarzania obrazu.....	41
5 Opis aplikacji.....	43
5.1 Postawione wymagania.....	43
5.2 Ogólny opis działania.....	43
5.3 Ustawienia poszczególnych procesów analizy obrazu.....	45
5.4 Obliczenia i prezentacja wyników.....	48
6 Testy systemu wizyjnego i ich wyniki.....	54
6.1 Analiza trajektorii odbijającej się piłki.....	55
6.2 Wpływ kalibracji i błędu analizy obrazu na obliczaną pozycję obiektu.....	59
6.3 Wpływ zmiany położenia aparatu.....	61
6.4 Czas przetwarzania obrazu.....	62

---

Wykorzystanie cienia do odtworzenia pozycji obiektu w przestrzeni

---

7 Kierunki dalszego rozwoju pracy.....	64
Podsumowanie.....	66
Literatura.....	68
Załączniki.....	69

## Indeks ilustracji

Rys. 1: Geometria epipolarna w systemie stereowizyjnym, źródło [1].....	11
Rys. 2: Układ kanoniczny systemu stereowizyjnego, źródło: [1].....	13
Rys. 3: Układ stereowizyjny wykorzystujący transformatę płaską do obliczenia pozycji obiektu.....	14
Rys. 4: Układ stereowizyjny wykorzystujący transformatę płaską oraz cień.....	15
Rys. 5: Akwizycja obrazu stereoskopowego w układzie katadioptrycznym, źródło: [1]	16
Rys. 6: Załamanie cienia na krawędzi sceny, źródło: [2].....	19
Rys. 7: Cień jako element obrazu niosący dodatkowe informacje, źródło: [2].....	20
Rys. 8: Przykładowe obrazy służące do kalibracji: wzorzec kalibracyjny oraz wykorzystanie znanych punktów sceny.....	23
Rys. 9: Obraz wejściowy i wynik transformacji płaskiej płaszczyzny podstawy.....	24
Rys. 10: Widok kalibracji punktu źródła światła - dwa rzucane cienie i ich przedłużenie prostymi w celu wyznaczenia punktu rzutu źródła światła na płaszczyznę podstawy....	25
Rys. 11: Wyznaczanie wysokości źródła światła.....	26
Rys. 12: Geometria systemu ze światłem punktowym (żarówka).....	29
Rys. 13: Schemat zastosowanego algorytmu.....	33
Rys. 14: Zmiany obrazu po kolejnych krokach algorytmu przetwarzania: a) usunięcie koloru, b) zastosowanie maski, zmniejszenie obszaru c) odjęcie tła, d) binaryzacja, e) erozja i oznaczenie obiektów, f) wyznaczenie środków obiektów.....	34
Rys. 15: Analizowany obraz przed i po zastosowaniu maski usuwającej zbędne obszary.....	36
Rys. 16: Odejmowanie tła: a) analizowany obraz, b) obraz tła, c) odejmowanie tła od obrazu, d) odejmowanie obrazu od tła, e) wartość bezwzględna odejmowania.....	38
Rys. 17: Obraz binarny i ten sam obraz po wykonaniu erozji i indeksowaniu obiektów	39
Rys. 18: Widok ogólny aplikacji – lista kroków przetwarzania obrazu oraz panel wczytywania obrazów.....	45
Rys. 19: Widok ustawień procesu skalowania.....	46
Rys. 20: Widok paneli ustawień: nakładania maski, odejmowania tła i binaryzacji.....	47
Rys. 21: Panele ustawień: erozja i indeksowanie oraz wyznaczenie środków obiektów	48

---

## Wykorzystanie cienia do odtworzenia pozycji obiektu w przestrzeni

---

Rys. 22: Panel kalibracji systemu.....	49
Rys. 23: Punkty kalibracji cienia.....	50
Rys. 24: Widok panelu obliczeń i prezentacji wyników.....	51
Rys. 25: Trajektoria ruchu odbijającej się piłeczki (światło słoneczne).....	56
Rys. 26: Trajektoria ruchu odbijającej się piłeczki (światło żarówki).....	57
Rys. 27: Wpływ zmiany położenia aparatu na wyznaczoną pozycję obiektu.....	62

## Wstęp

Ciągły rozwój komputerów i ich rosnąca moc obliczeniowa pozwalają na coraz szersze zastosowanie ich w przemyśle. Jedną ze stale rozwijanych gałęzi rozwiązań przemysłowych są systemy wizyjne. Próby naśladowania zdolności widzenia człowieka pozwalają na pozyskanie z obrazów nowych informacji i zautomatyzowanie kolejnych procesów produkcyjnych.

Cień występujący na obrazie dawniej rozważany był jedynie w kategoriach zakłócenia. Wciąż aktualne są publikacje na temat jego usuwania z obrazu, przede wszystkim w przypadku systemów montowanych na zewnątrz, np. obserwujących ruch uliczny. Okazuje się jednak, że nie zawsze jest on tylko elementem utrudniającym analizę. Cień może również dostarczyć dodatkowych informacji, których nie można uzyskać obserwując jedynie obiekty rzeczywiście występujące na obrazie. Dostarczone dane mogą odnosić się do wskazania obecności obiektów, które nie są bezpośrednio widoczne na obrazie – są przysłonięte lub znajdują się w pobliżu sceny. Możliwe do uzyskania są również dane o źródle światła, a nawet mapie głębi otoczenia.

W niniejszej pracy analizowane jest wyznaczanie pozycji obiektu w przestrzeni wykorzystując obecny na obrazie cień. Systemy wizyjne określające położenie obiektu zostały już dobrze opisane i są wykorzystywane w przemyśle. Zazwyczaj są to dwie kamery rejestrujące dwa obrazy tej samej sceny z różnej perspektywy. Takie systemy stereowizyjne pozwalają na stosunkowo dokładne określenie pozycji obiektu w przestrzeni. Okazuje się jednak, że możliwe jest ograniczenie ilości potrzebnych obrazów z dwóch do jednego. Jeden obraz, na którym znajduje się obiekt i jego cień, zawiera wystarczającą ilość danych do wyznaczenia położenia. Konceptyjnie, jedna z kamer musi w takim przypadku zostać zastąpiona przez źródło światła i może być to zarówno światło punktowe, np. żarówka, jak i słońce, które z punktu widzenia obserwatora na Ziemi emituje promienie równoległe. Zmienia się wówczas geometria systemu, lecz zachowana zostaje jego funkcjonalność.



W systemie wizyjnym wyznaczającym pozycję obiektu w przestrzeni użycie jednej kamery zamiast dwóch, niesie za sobą pewne korzyści. Rozwiązanie takie redukuje koszty wprowadzenia systemu wizyjnego na linii produkcyjnej, jak również zmniejsza awaryjność systemu. Miejsce montażu źródła światła jest również łatwiejsze i bardziej elastyczne niż kamery, ze względu na ograniczenia techniczne takie jak odpowiednia temperatura pracy, czy występujące drgania. Cena żarówki pozwala także na montaż jej w miejscach, w których jest narażona na uszkodzenie.

Wykorzystanie zaproponowanego algorytmu nie musi się ograniczać do budowy całego systemu wizyjnego. Do wyznaczenia pozycji obiektu w przestrzeni wystarczającym okazuje się dowolne zdjęcie, na którym można przeprowadzić proces kalibracji poprzez wskazanie odpowiedniej ilości znanych punktów. Możliwa jest zatem analiza położenia obiektów na istniejących już zdjęciach, niezarejestrowanych przy użyciu metod stereoskopowych.

### ***Cel i zakres pracy***

Celem pracy jest stworzenie algorytmu pozwalającego na wyznaczenie pozycji obiektu w przestrzeni na podstawie jednego obrazu na którym znajduje się obiekt i jego cień oraz implementacja odpowiedniej aplikacji pozwalającej na weryfikację zaproponowanej metody. Ostatnim elementem jest wykonanie serii zdjęć, przeprowadzenie testów algorytmu i aplikacji oraz ich ocena.

### ***Struktura pracy***

W rozdziale 1. przedstawiono tradycyjne podejście do stereowizji, czyli różne modele wykorzystywanej geometrii takich systemów. W rozdziale 2. przytoczone są informacje o znaczeniu cienia w komputerowej analizie obrazów. Rozdział 3. zawiera zaproponowaną metodę wyznaczania pozycji obiektu w przestrzeni na podstawie cienia. Znajdują się w nim trzy podrozdziały, mówiące o kalibracji płaszczyzn wykorzystywanych do analizy (3.1), kalibracji światła (3.2) oraz samo wyznaczenie pozycji obiektu w przestrzeni (3.3). Rozdział 4. zawiera

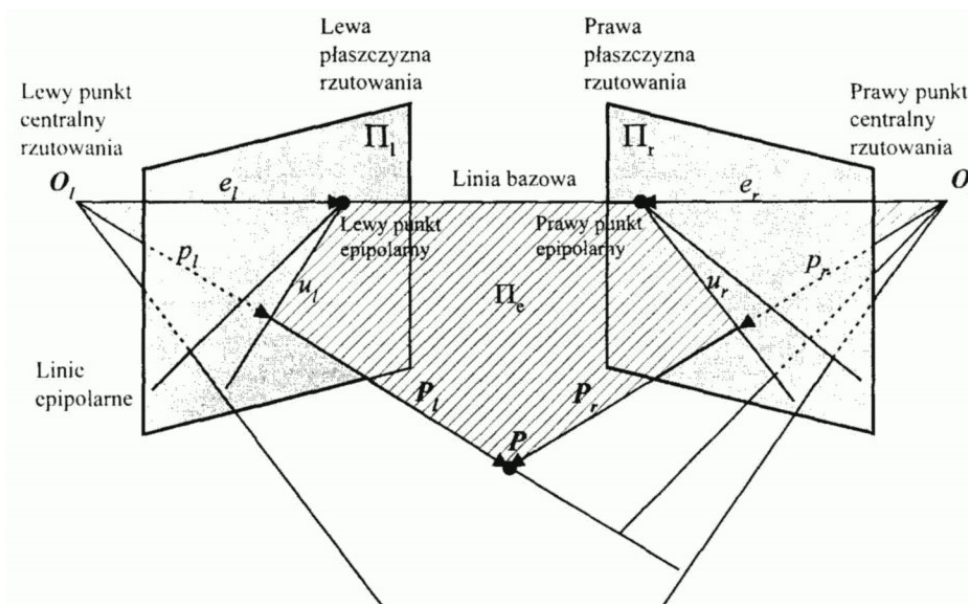
opis zastosowanych algorytmów przetwarzania obrazu, używanych w celu wyznaczenia położenia obiektu i cienia na obrazie. W rozdziale 5. przedstawiono stworzoną aplikację. Wymieniono w nim najważniejsze postawione wobec niej wymagania oraz opisano ogólny schemat jej działania, możliwości ustawień procesów przetwarzania i sposób prezentacji wyników obliczeń. Przeprowadzone testy zostały omówione w rozdziale 6. Dotyczą one przede wszystkim powtarzalności i dokładności wyznaczonej pozycji. Rozdział 7. zawiera informacje o możliwościach rozwoju pracy i aplikacji. Całość została zakończona podsumowaniem zawierającym ocenę zaproponowanej metody wyznaczenia położenia obiektu w przestrzeni.

## 1 Tradycyjna stereowizja

Rozpoznawanie głębi jest naturalną częścią procesu widzenia człowieka. Na podstawie obrazów z obu oczu, potrafi on określić wzajemne położenie widzianych obiektów oraz ocenić w przybliżeniu w jakiej odległości się znajdują. Współczesna technika stara się jedynie naśladować niektóre zachowania natury. Układ wzrokowy stał się podstawą rozważań o stereowizji i prób budowy systemów wyznaczających głębię sceny. Pomimo szybko rozwijającej się nauki, oko ludzkie wciąż jest dużo bardziej zaawansowanym narzędziem w rozpoznawaniu obrazu niż jakiegokolwiek urządzenie wymyślone przez człowieka.

### 1.1 Geometria epipolarna

Na Rys. 1 przedstawiono podstawowy układ kamer systemu stereowizyjnego obserwujących punkt  $P$ . Płaszczyzna  $\Pi_l$  oraz  $\Pi_r$ , wraz z odpowiadającymi im punktami rzutowania  $O_l$  i  $O_r$ , stanowią model aparatu o obiektywie punktowym.



Rys. 1: Geometria epipolarna w systemie stereowizyjnym, źródło [1]

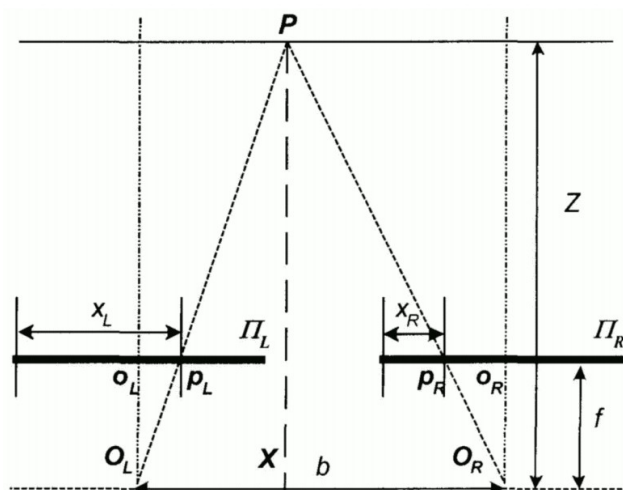
Proste  $\overline{PO}_l$  i  $\overline{PO}_r$  wyznaczają obrazy punktu  $P$  na płaszczyznach  $\Pi_l$  oraz  $\Pi_r$  (odpowiednio  $p_l$  i  $p_r$ ), które po dyskretyzacji odpowiadają współrzędnym pikselowym szukanego punktu na zarejestrowanym obrazie cyfrowym. Znając zatem punkty  $p_l$  i  $p_r$  oraz mając odpowiednio skalibrowany system, możemy wyznaczyć przybliżenia prostych  $\overline{PO}_l$  i  $\overline{PO}_r$ . Ze względu na niedokładności kalibracji, obliczeń i występującą dyskretyzację, proste te w ogólnym przypadku nie będą się przecinać. Przybliżenie punktu  $P$  jest wyznaczane więc jako środek najkrótszego odcinka pomiędzy tymi prostymi.

Wspomniana powyżej kalibracja jest jednym z ważniejszych elementów budowy systemu wizyjnego, gdyż od jej dokładności zależeć będzie dokładność całego systemu. Służy ona wyznaczeniu współczynników przekształcenia obserwowanej płaszczyzny w przestrzeni trójwymiarowej do dwuwymiarowej przestrzeni rejestrowanego obrazu. Współczynniki te odpowiadają wewnętrznym i zewnętrznym parametrom kamery, takim jak położenie, obrót czy ogniskowa.

Geometria epipolarna posiada dodatkową cechę ułatwiającą przyporządkowanie odpowiadających sobie punktów na obu obrazach. Linia bazowa, która łączy punkty rzutowania  $O_l$  i  $O_r$  przecina płaszczyzny  $\Pi_l$  i  $\Pi_r$ , tworząc punkty epipolarne, odpowiednio  $e_l$  i  $e_r$ , natomiast wraz z punktem  $P$  wyznacza ona płaszczyznę epipolarną  $\Pi_e$ . Punkty epipolarne, wraz z odpowiednimi punktami  $p_l$  i  $p_r$ , wyznaczają na płaszczyznach rzutowania tak zwane linie epipolarne. Linie te mogą być również wyznaczone jako proste powstałe z przecięcia płaszczyzn  $\Pi_l$  i  $\Pi_r$  z płaszczyzną epipolarną. Okazuje się, że punkty  $p_l$  i  $p_r$  odpowiadające szukanemu punktowi  $P$  mogą leżeć jedynie na liniach epipolarnych. Posiadając zatem informacje o punktach  $p_l$  i  $e_l$  na lewym obrazie, można wyznaczyć linię epipolarną na obrazie prawym, gdzie musi znajdować się punkt  $p_r$ . Upraszcza to proces dopasowania punktów – przeszukiwanie dwuwymiarowe zostaje zastąpione jednowymiarowym.

Szczególnym przypadkiem układu stereowizyjnego, jest układ kanoniczny,

zwany również układem standardowym lub podstawowym. Został on przedstawiony na Rys. 2. W tym przypadku osie optyczne obu kamer są równoległe, a punkty epipolarne przesunięte są do nieskończoności – linia bazowa łącząca środki rzutowania nie przecina płaszczyzn rzutowania. Pozwala to dodatkowo uprościć proces dopasowania punktów pomiędzy obrazami. Linie poszukiwania odpowiadających punktów na obu poszczególnych płaszczyznach są współliniowe i poziome.



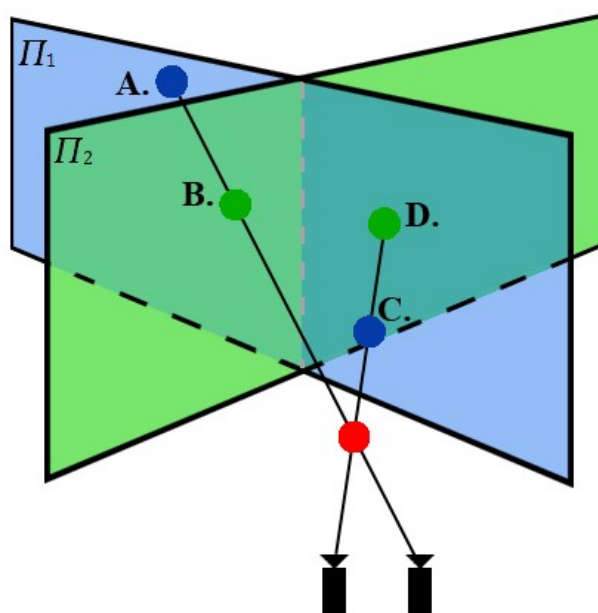
Rys. 2: Układ kanoniczny systemu stereowizyjnego, źródło: [1]

Istnieje możliwość przekształcenia obrazów stereoskopowych w sposób umożliwiający taką ich analizę, jakby były zarejestrowane w układzie kanonicznym. Przekształcenie to nazwane jest rektyfikacją.

## 1.2 System stereowizyjny oparty na transformacji płaskiej

Innym podejściem do analizy obrazów stereoskopowych jest wykorzystanie transformacji płaskiej. Pozwala ona znaleźć przekształcenie, które występuje podczas rzutowania dowolnie ułożonej w przestrzeni płaszczyzny na płaszczyznę obrazu kamery oraz przekształcenie do niego odwrotne. Sposób wyznaczenia

tej transformacji został szczegółowo opisany w podrozdziale 3.1. Na Rys. 3 przedstawiono schemat układu stereowizyjnego wykorzystującego tę transformację. Szukany punkt  $P$  widoczny jest przez obie kamery, a punkty  $A, B, C$  i  $D$  są punktami na płaszczyznach  $\Pi_1$  i  $\Pi_2$ , które z punktu widzenia kamer są rzutami punktu  $P$  na te płaszczyzny.



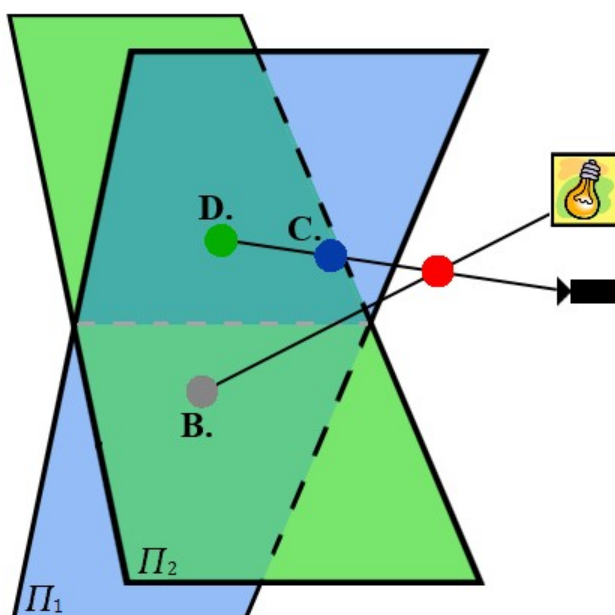
Rys. 3: Układ stereowizyjny wykorzystujący transformatę płaską do obliczenia pozycji obiektu

Transformata płaska pozwala na wyznaczenie współrzędnych poszczególnych punktów  $A, B, C$  i  $D$  na odpowiednich płaszczyznach w przestrzeni, co przy znajomości położenia płaszczyzn określa jednoznacznie przestrzenne współrzędne tych punktów. Kolejnym krokiem jest zatem znalezienie prostych  $\overline{AB}$  i  $\overline{CD}$ . Punkt przecięcia tych dwóch prostych, lub dokładniej środek najkrótszego odcinka łączącego te proste, jest przybliżeniem szukanego punktu.

### 1.3 Systemy stereowizyjne z jedną kamerą

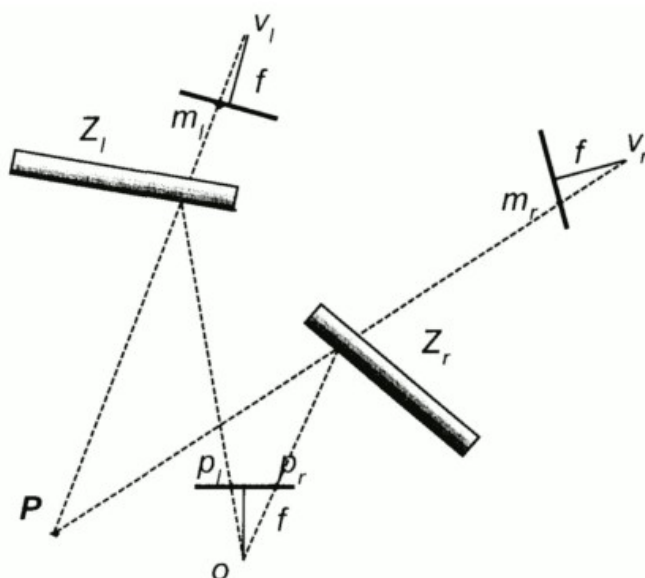
Układ opisany powyżej, oparty na geometrii wykorzystującej transformatę

płaską stanowił podstawę do stworzenia systemu wizyjnego wykorzystującego tylko jedną kamerę oraz cień rzucany przez szukany obiekt. Rys. 4 przedstawia zmodyfikowany system z Rys. 3 po uwzględnieniu zamiany jednej kamery na źródło światła, w tym przypadku punktowe – żarówkę. Działanie to zmienia strukturę systemu. Zastosowanie transformaty płaskiej umożliwia wyznaczenie dwóch punktów –  $C$  i  $D$ , powstałych poprzez rzutowanie punktu  $P$  na płaszczyzny  $\Pi_1$  i  $\Pi_2$ . Pozwala ona również na wyznaczenie pozycji cienia, gdyż leży on na jednej ze skalibrowanych płaszczyzn. Możliwe jest zatem wyznaczenie prostej  $\overline{AB}$ , która przebiega przez poszukiwany obiekt. Druga prosta może być natomiast wyznaczona jako prosta łącząca cień ze źródłem światła. Dwie różne proste pozwalają, tak jak w poprzednio analizowanym systemie, na wyznaczenie przybliżonej pozycji poszukiwanego obiektu w przestrzeni.



Rys. 4: Układ stereowizyjny wykorzystujący transformatę płaską oraz cień

Wykorzystanie jednej kamery do zarejestrowania obrazów stereoskopowych może odbywać się również w układzie katadioptrycznym. Zastosowanie elementów odbijających światło pozwala na zwiększenie pola widzenia kamer. Odpowiednie ułożenie zwierciadeł pozwala również na budowę stereowizyjnego systemu, co pokazane zostało na Rys. 5.



Rys. 5: Akwizycja obrazu stereoskopowego w układzie katadioptrycznym, źródło: [1]

Środek rzutowania obrazu, czyli ognisko rzeczywistej kamery, oznaczono literą  $o$ . Poprzez odbicie obrazu obserwowanego punktu  $P$  w zwierciadłach  $Z_l$  oraz  $Z_r$  powstają pozorne środki rzutowania, oznaczone  $v_l$ ,  $v_r$ , które są odpowiednikami dwóch osobnych kamer. Na zarejestrowanym obrazie powstają ostatecznie dwa obrazy punktu  $P$ , co przy znajomości parametrów wewnętrznych i zewnętrznych aparatu oraz zwierciadeł pozwala na określenie położenia punktu  $P$  w przestrzeni.

Układ katadioptryczny i zaproponowany układ z wykorzystaniem cienia posiadają szereg zalet w porównaniu do klasycznych systemów stereowizyjnych



wykorzystujących dwie kamery:

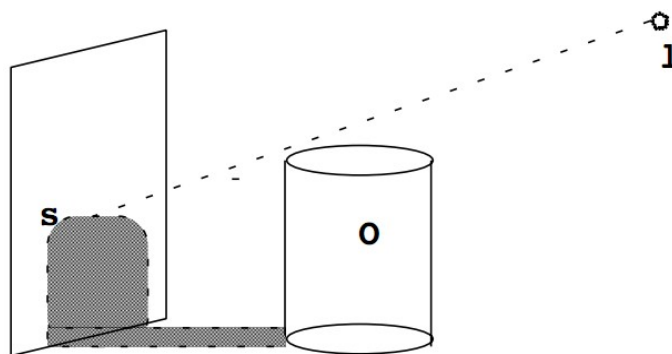
- łatwiejsza kalibracja – określenie parametrów tylko dla jednego aparatu,
- brak konieczności synchronizacji aparatów – np. wyzwalenie migawki,
- zmniejszenie kosztów systemu wizyjnego,
- możliwość zwiększenia pola widzenia poprzez zastosowanie zwierciadeł parabolicznych w układzie katadioptrycznym.

Bez względu na wykorzystany układ kamer, a nawet ich ilość, do odtworzenia pozycji obiektu w przestrzeni zawsze potrzebne są co najmniej dwa jego obrazy. Możliwe jest jednak wykorzystanie jednej kamery do zarejestrowania obu potrzebnych obrazów.

## 2 Cień w systemach wizyjnych

Występujący na analizowanych obrazach cień zazwyczaj uważany jest za zakłócenie. Powoduje on zniekształcenie obserwowanych obiektów, a także może zostać błędnie oznaczony jako jeden z nich, np. w przypadku użycia procesu odejmowania tła. Cień w systemach wizyjnych jest zatem zjawiskiem, które próbuje się eliminować. Na liniach produkcyjnych możliwe jest zastosowanie odpowiedniego oświetlenia lub nawet namiotów bezcieniowych używanych przez fotografów. Tworzone są również algorytmy, które mają za zadanie usunąć cienie z obrazu.

Okazuje się jednak, że występujący cień niesie za sobą również dużo dodatkowych danych. Występuje tutaj ścisła zależność pomiędzy źródłem światła, cieniem oraz obiektem który go rzuca [2]. Istnieje kilka możliwości uzyskania różnych informacji z obrazu. Pierwszą z nich jest położenie obiektu w przestrzeni przy znajomości położenia źródła światła oraz płaszczyzny (ekranu), na który rzucony jest cień (dokładny opis metody został zaprezentowany w podrozdziale 3.3). Możliwe jest również uzyskanie informacji o pozycji źródła światła przy znajomości położenia ekranu i obiektów rzucających cień. To zjawisko również zostało zastosowane w niniejszej pracy do kalibracji źródła światła, czyli jego położenia przy świetle punktowym lub kierunku padania dla światła słonecznego – podrozdział 3.2. Możliwa jest również ocena intensywności źródła światła. Trzecią analogiczną sytuacją jest uzyskanie informacji o ekranie, na który rzucony jest cień. Przykładowo, użycie prostego, pionowego elementu rzucającego cień pozwala obserwować załamania linii cienia na krawędziach obiektów występujących na scenie, co przedstawia Rys. 6. Zjawisko to można wykorzystać do zbudowania mapy głębi obserwowanej sceny.



Rys. 6: Załamanie cienia na krawędzi sceny, źródło: [2]

Położenie i kształt cienia pozwala więc na uzyskanie nowych informacji z obrazu. Samo występowanie cienia może również dostarczyć kolejnych danych. Pojawia się tutaj trudność z wykorzystaniem takich zjawisk w systemach wizyjnych, ponieważ dane które niesie cień są często nieprecyzyjne. Przykład został zaprezentowany na Rys. 7. Na lewej fotografii widoczny jest jedynie pusty kamienny blok. Cień, który się na nim znajduje pozwala jednak wnioskować, że zaraz obok znajduje się człowiek. Podobnie na prawym obrazie widoczna jest jedna donica, ale cień wskazuje, że za nią znajduje się inny obiekt. Informacja uzyskana w ten sposób jest jednak nieprecyzyjna. Nie jest możliwe dokładne określenie co znajduje się za donicą. Cień z lewej fotografii niekoniecznie jest również rzucony przez człowieka. Pozyskana jest informacja jedynie o obecności dodatkowego obiektu w obrębie lub otoczeniu sceny pomimo tego, że nie jest on bezpośrednio obserwowany przez kamerę.



*Rys. 7: Cień jako element obrazu niosący dodatkowe informacje, źródło: [2]*

### 3 Geometria systemu stereowizyjnego wykorzystującego cień obiektu

#### 3.1 Kalibracja płaszczyzn - transformacja „płaska”

Pierwszym elementem umożliwiającym analizę geometrii uchwyconej na obrazie sceny jest znalezienie przyporządkowania punktów płaszczyzny obrazu do punktów w przestrzeni. W tym celu użyta została transformacja płaska [3]. Przyporządkowanie to jest oparte na przekształceniu jednokładności pomiędzy płaszczyzną obrazu (czyli matrycy CCD aparatu, ekranu) i jedną, wybraną płaszczyzną w przestrzeni, która jest identyfikowana.

Ogólna postać transformacji płaskiej jest następująca:

$$\begin{bmatrix} x_{P_1} \\ y_{P_1} \end{bmatrix} = \frac{1}{\alpha x_{P_2} + \beta y_{P_2} + 1} \cdot \begin{bmatrix} a_1 x_{P_2} + b_1 y_{P_2} + c_1 \\ a_2 x_{P_2} + b_2 y_{P_2} + c_2 \end{bmatrix}$$

gdzie:

$$\begin{bmatrix} x_{P_1} \\ y_{P_1} \end{bmatrix} - \text{współrzędne punktu na płaszczyźnie 1}$$

$$\begin{bmatrix} x_{P_2} \\ y_{P_2} \end{bmatrix} - \text{współrzędne punktu na płaszczyźnie 2}$$

$a_1, b_1, c_1, a_2, b_2, c_2, \alpha, \beta$  - współczynniki transformacji

Współczynniki transformacji odzwierciedlają wewnętrzne oraz zewnętrzne parametry kamery, takie jak ogniskowa, sposób montażu soczewki i przetwornika CCD, wzajemne położenie kamery względem badanej płaszczyzny itp. Wprowadzone zostaje oznaczenie wektorowe, pozwalające na agregację wszystkich współczynników:

$$H = [a_1 \quad b_1 \quad c_1 \quad a_2 \quad b_2 \quad c_2 \quad \alpha \quad \beta]^T$$

Do znalezienia współczynników transformacji wystarczą jedynie cztery punkty znajdujące się na płaszczyźnie obrazu oraz odpowiadające im punkty na identyfikowanej płaszczyźnie w przestrzeni. Możliwe jest również zastosowanie tej metody używając większej ilości zestawów punktów, co podwyższa dokładność kalibracji. W celu wyznaczenia współczynników transformacji należy rozwiązać równanie liniowe:

$$\begin{bmatrix} x_{r1} \\ y_{r1} \\ \vdots \\ x_{rn} \\ y_{rn} \end{bmatrix} = \begin{bmatrix} x_{m1} & y_{m1} & 1 & 0 & 0 & 0 & -x_{m1}x_{r1} & -y_{m1}x_{r1} \\ 0 & 0 & 0 & x_{m1} & y_{m1} & 1 & -x_{m1}y_{r1} & -y_{m1}y_{r1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{mn} & y_{mn} & 1 & 0 & 0 & 0 & -x_{mn}x_{rn} & -y_{mn}x_{rn} \\ 0 & 0 & 0 & x_{mn} & y_{mn} & 1 & -x_{mn}y_{rn} & -y_{mn}y_{rn} \end{bmatrix} H$$

gdzie:

$x_{r,1}, y_{r,1}, \dots, x_{r,n}, y_{r,n}$  - rzeczywiste współrzędne kolejnych punktów kalibracji

$x_{m,1}, y_{m,1}, \dots, x_{m,n}, y_{m,n}$  - współrzędne obrazowe kolejnych punktów kalibracji

$H$  - szukany wektor współczynników transformacji

Kalibracja jednej płaszczyzny polega zatem na wskazaniu punktów na obrazie i przyporządkowaniu im współrzędnych płaszczyznowych w rzeczywistej przestrzeni.

Aby możliwe było wyznaczenie współrzędnych punktu w przestrzeni trójwymiarowej, należy odnaleźć przekształcenia transformacji dla dwóch różnych płaszczyzn. Przyjęto oznaczenia: płaszczyzna podstawy, na której rejestrowany jest cień obiektu, oraz płaszczyzna pionowa.

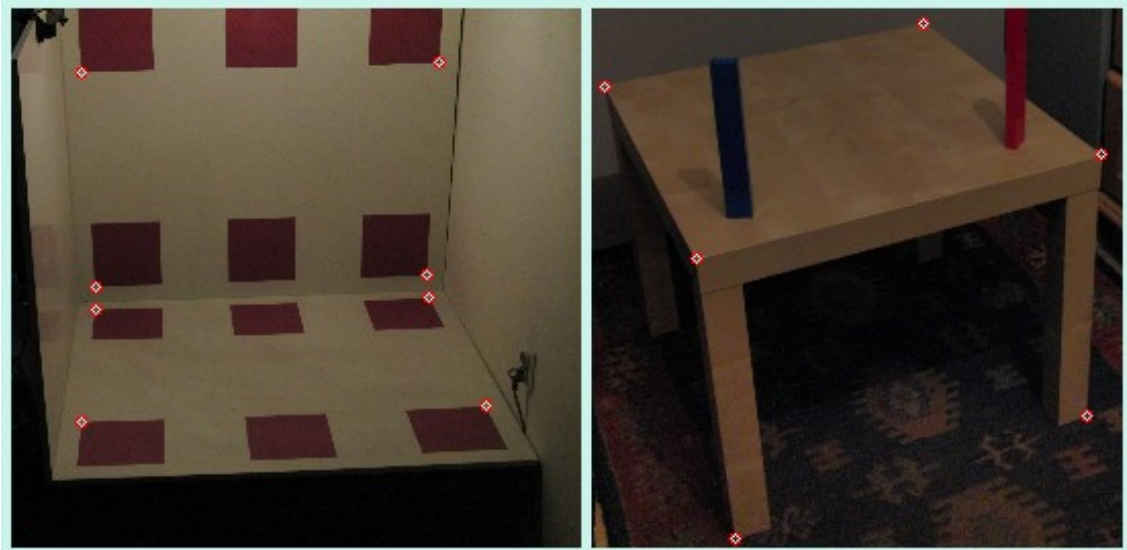
Kalibracja przeprowadzana jest na podstawie wzorca kalibracyjnego, którego wymiary są znane – przykładowy został zaprezentowany na Rys. 8. Widoczne są na nim kwadraty umieszczone na płaszczyźnie podstawy i płaszczyźnie pionowej. Do kalibracji mogą posłużyć zarówno środki kwadratów

---

## Wykorzystanie cienia do odtworzenia pozycji obiektu w przestrzeni

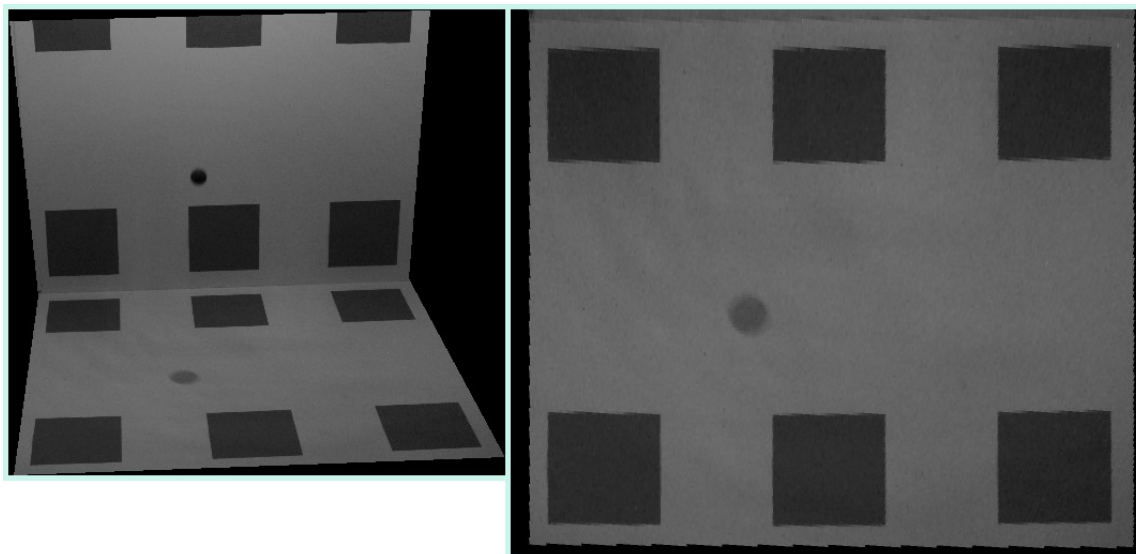
---

jak i ich wierzchołki. Na drugim obrazie zaprezentowano również inne podejście - wykorzystanie naturalnie występujących płaszczyzn otoczenia rejestrowanej sceny i wprowadzenie punktów o znanym położeniu. Pozwala to uniknąć budowy wzorca i odpowiedniego ustawiania go względem otoczenia.



*Rys. 8: Przykładowe obrazy służące do kalibracji: wzorec kalibracyjny oraz wykorzystanie znanych punktów sceny*

Transformacja płaska umożliwia również odtworzenie obrazu identyfikowanych płaszczyzn w widoku „z góry”, a brak zniekształceń widocznych po jego wyświetleniu pozwala potwierdzić poprawność przeprowadzonej kalibracji (Rys. 9).



Rys. 9: Obraz wejściowy i wynik transformacji płaskiej płaszczyzny podstawy

### **3.2 Kalibracja światła – wyznaczenie współrzędnych punktu światła lub jego kąta padania**

Kolejnym elementem, który musi zostać skalibrowany przed właściwym uruchomieniem systemu jest światło. Źródła światła, dla potrzeb niniejszej pracy, należy podzielić ze względu na geometrię jego rozchodzenia się:

- światło rozchodzące się promieniście (żarówka),
- światło o promieniach biegnących równoległe (słońce).

Ze względu na dużą odległość Ziemi od Słońca można założyć, że promienie słoneczne w jej pobliżu rozchodzą się równoległe, podczas gdy żarówka emituje światło, które rozchodzi się promieniście. Różnica ta wprowadza konieczność innej kalibracji światła oraz odmiennej analizy geometrycznej podczas obliczania pozycji obiektu.

#### **3.2.1 Źródło światła o promieniach rozchodzących się promieniście – żarówka**

Kalibracja światła, w przypadku użyciu żarówki, sprowadza się



---

### Wykorzystanie cienia do odtworzenia pozycji obiektu w przestrzeni

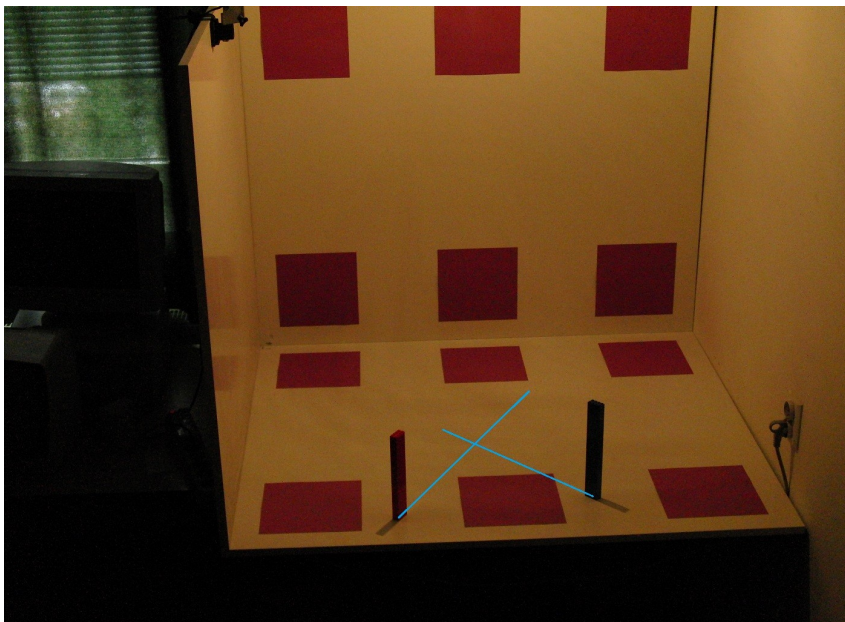
---

do wyznaczenia punktu jej położenia w przestrzeni. Procedura składa się z dwóch elementów (konieczne jest wcześniejsze skalibrowanie płaszczyzny podstawy):

- wyznaczenie rzutu prostopadłego punktu źródła światła na płaszczyznę podstawy,
- wyznaczenie wysokości na jakiej znajduje się źródło światła.

Do kalibracji potrzebne jest wykonanie zdjęcia, gdy w obrębie rejestrowanej sceny znajdują się dwa pionowo ustawione bloczki o znanej wysokości. Wymagane jest również aby widoczny był ich cień w taki sposób, aby można było oznaczyć jego początek i koniec (punkt u podstawy obiektu oraz punkt wierzchołka cienia).

Po oznaczeniu powyżej opisanych punktów wyznaczone zostają proste, wzdłuż których leżą cienie obu obiektów, co przedstawia Rys. 10. Punkt przecięcia tych prostych jest punktem rzutu prostopadłego źródła światła na płaszczyznę podstawy. Współrzędne tego punktu na płaszczyźnie podstawy, są zarazem współzrędnymi  $X$  i  $Y$  punktu światła.



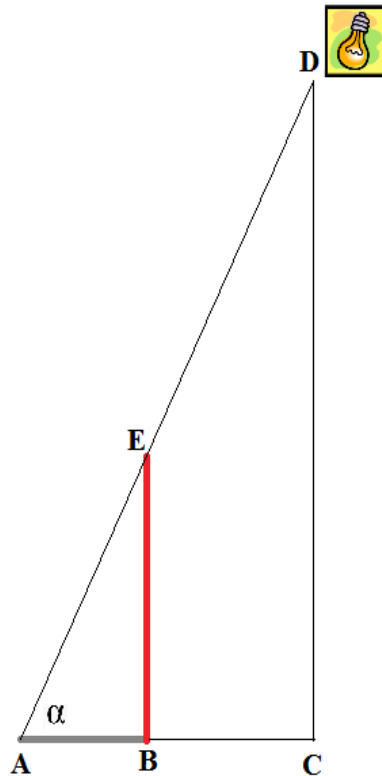
*Rys. 10: Widok kalibracji punktu źródła światła - dwa rzucane cienie i ich przedłużenie prostymi w celu wyznaczenia punktu rzutu źródła światła na płaszczyznę podstawy*

---

## Wykorzystanie cienia do odtworzenia pozycji obiektu w przestrzeni

---

Następnie należy wyznaczyć wysokość położenia punktu światła. Ze względu na ten krok oznaczane wcześniej punkty nie mogły być wybrane dowolnie na prostej cienia, lecz powinny wskazywać na jego początek i koniec. Rys. 11 przedstawia koncepcję obliczeń w oparciu o twierdzenie Talesa.



Rys. 11: Wyznaczanie wysokości źródła światła

Przyjęte oznaczenia:

punkt  $A$  – punkt końca cienia

punkt  $B$  – punkt początku cienia, czyli punkt podstawy bloczka kalibracyjnego

punkt  $C$  – punkt rzutu źródła światła na płaszczyznę

punkt  $D$  – punkt źródła światła

punkt  $E$  – punkt szczytu bloczka kalibracyjnego (rzucający cień na punkt  $A$ )

$|BE|$  – wysokość bloczka kalibracyjnego – wartość zmierzona ręcznie

Obliczenie wysokości na jakiej znajduje się źródło światła wynika z twierdzenia Talesa (lub podobieństwa trójkątów):

$$|CD| = \frac{|AC| \cdot |BE|}{|AB|}$$

W celu zminimalizowania błędu spowodowanego niedokładnym oznaczeniem punktów na obrazie, wysokość położenia źródła światła obliczana jest osobno dla obu bloczków, a następnie wykorzystywana jest wartość średnia.

Posiadając współrzędne źródła światła oraz punkt cienia na płaszczyźnie podstawy, możliwe jest wyznaczenie prostej przechodzącej przez te punkty, co jest elementem obliczania położenia obiektu w przestrzeni.

### **3.2.2 Źródło światła o promieniach równoległych – słońce**

Kalibracja źródła światła, gdy wykorzystywane jest słońce, różni się od tej, którą stosuje się w przypadku użycia żarówki. Kąt padania światła z żarówki jest zmienny – zmienia się w zależności od położenia obiektu. Stały jest jednak punkt źródła światła, więc prosta przechodząca przez cień, obiekt i źródło światła może zostać wyznaczona na podstawie punktu położenia światła i punktu cienia. Gdy wykorzystywane jest światło słoneczne, to wektor padania promieni jest stały (w danym, dostatecznie krótkim czasie) i dlatego kalibracja wyznacza tę informację.

Do kalibracji światła potrzebny jest w tym przypadku tylko jeden bloczek kalibracyjny. Wektor padania światła obliczany jest jako różnica punktu na wierzchołku cienia i punktu na szczycie bloczka kalibracyjnego. Prosta przechodząca przez punkt cienia, obiekt i źródło światła (słońce) jest w tym przypadku wyznaczona jednoznacznie przez punkt cienia i wektor padania promieni światła.

Zmiana wektora padania promieni słonecznych w czasie, spowodowana jest ruchem obrotowym Ziemi. Dla dostatecznie krótkich odcinków czasu efekt ten może być pominięty lecz w przypadku analizy dłuższej serii zdjęć konieczna jest

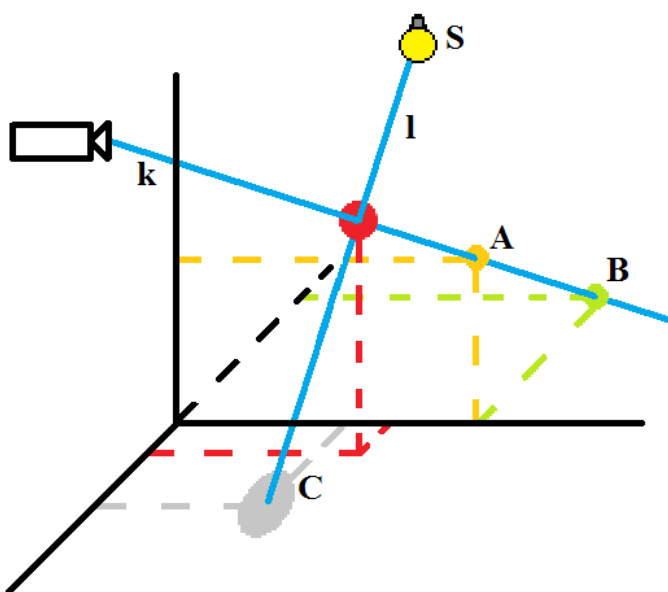
korekta wyznaczonego wektora. Może być to wykonane poprzez cykliczną (zautomatyzowaną) kalibrację światła lub poprzez obliczanie kąta padania światła na podstawie czasu oraz informacji o położeniu i orientacji obserwowanej sceny (np. dane z systemu GPS).

### **3.3 Wyznaczanie położenia obiektu**

Wyznaczenie współczynników transformacji płaskich, poprzez kalibrację opisaną w poprzednich podrozdziałach, pozwala, przy użyciu odpowiedniego przekształcenia, przypisać poszczególnym punktom na obrazie współrzędne ich rzutów na odpowiednie płaszczyzny w rzeczywistym układzie.

Po odnalezieniu na obrazie punktów, w których znajdują się środki obiektu i jego cienia, należy znaleźć współrzędne ich rzutów na skalibrowane płaszczyzny. Punkt cienia rzutowany jest jedynie na płaszczyznę podstawy. Ponieważ znajduje się on na tej płaszczyźnie, znane są jego rzeczywiste współrzędne w przestrzeni. Kolejnym krokiem jest rzutowanie punktu środka obiektu na obie płaszczyzny – pionową oraz podstawy.

Rys. 12 przedstawia położenie omówionych punktów oraz koncepcję dalszych obliczeń. Rzutowany punkt cienia oznaczony został literą  $C$ , podczas gdy litery  $A$  oraz  $B$  odpowiadają rzutom punktu środka obiektu na płaszczyznę pionową oraz płaszczyznę podstawy. Kolejnym etapem jest przeprowadzenie dwóch prostych przez dwie pary punktów. Pierwsza prosta przeprowadzona przez punkty  $A$  i  $B$  składa się ze wszystkich tych punktów, w których hipotetycznie mogłyby znajdować się obiekt, bez zmiany swojego położenia na zarejestrowanym obrazie. Druga prosta wyznaczana jest na dwa sposoby w zależności od użytego źródła światła. Przy świetle żarówki, prosta przechodzi przez punkt cienia oraz przez skalibrowany punkt źródła światła, natomiast przy świetle słonecznym prosta wyznaczana jest na podstawie punktu cienia oraz wektora kierunku padania światła.



Rys. 12: Geometria systemu ze światłem punktowym (żarówka)

Posiadając dwie różne proste, na których znajduje się obiekt, możliwe jest wyznaczenie jego współrzędnych w przestrzeni. W idealnym przypadku należałoby znaleźć punkt przecięcia obu prostych. W rzeczywistości można założyć, że proste te nie mają punktu wspólnego. Jest to spowodowane przede wszystkim błędami kalibracji, błędem wyznaczenia środków obiektu i cienia, oraz dyskretyzacją, jaka jest związana z cyfrowym zapisem obrazu. Szukany jest zatem najkrótszy odcinek łączący obie proste. Za punkt, w którym znajduje się obiekt, przyjmowany jest jego środek. W szczególnym przypadku odcinek ten może być zerowej długości, a jego końcowe punkty mogą być jednakowe. Nie zaburza to jednak obliczeń poszukiwania tego odcinka ani szukania jego środka. Uwzględniona jest zatem sytuacja, gdy proste rzeczywiście się przecinają.

Obliczanie szukanego punktu, rozpoczyna się od przedstawienia prostych w zapisie wektorowym:

$$\begin{aligned} k &: A + m \cdot \vec{AB} \\ l &: C + m \cdot \vec{CS} \end{aligned}$$

gdzie:

$$A, B, C, S, \vec{AB}, \vec{CS} \in \mathbb{R}^3$$

$$m, n \in \mathbb{R}$$

Następnie funkcja, która określa długość dowolnego odcinka rozpiętego pomiędzy prostymi  $k$  i  $l$ , jest minimalizowana.

$$F(m, n) = \|A + m \cdot \vec{AB}, C + n \cdot \vec{CS}\| \rightarrow \min$$

$$F(m, n) = [(A_x + m \cdot (B_x - A_x) - C_x - n \cdot (S_x - C_x))^2 + (A_y + m \cdot (B_y - A_y) - C_y - n \cdot (S_y - C_y))^2 + (A_z + m \cdot (B_z - A_z) - C_z - n \cdot (S_z - C_z))^2]^{-\frac{1}{2}} \rightarrow \min$$

Ekstremum funkcji znajdować się może jedynie w miejscu, gdzie zeruje się gradient. Wynosi on zatem:

$$\nabla F = \begin{bmatrix} \frac{\partial F}{\partial m} \\ \frac{\partial F}{\partial n} \end{bmatrix}$$

gdzie:

$$\frac{\partial F}{\partial m} = -\frac{1}{2} \cdot 2(A_x + m(B_x - A_x) - C_x - n(D_x - C_x))(B_x - A_x)$$

$$-\frac{1}{2} \cdot 2(A_y + m(B_y - A_y) - C_y - n(D_y - C_y))(B_y - A_y)$$

$$-\frac{1}{2} \cdot 2(A_z + m(B_z - A_z) - C_z - n(D_z - C_z))(B_z - A_z)$$

$$\frac{\partial F}{\partial n} = \frac{1}{2} \cdot 2(A_x + m(B_x - A_x) - C_x - n(D_x - C_x))(D_x - C_x)$$

$$+\frac{1}{2} \cdot 2(A_y + m(B_y - A_y) - C_y - n(D_y - C_y))(D_y - C_y)$$

$$+\frac{1}{2} \cdot 2(A_z + m(B_z - A_z) - C_z - n(D_z - C_z))(D_z - C_z)$$

---

## Wykorzystanie cienia do odtworzenia pozycji obiektu w przestrzeni

---

Przyrównanie gradientu funkcji do zera pozwala, po przekształceniach, otrzymać następujący liniowy układ równań:

$$\begin{bmatrix} a & -b \\ b & c \end{bmatrix} \begin{bmatrix} m \\ n \end{bmatrix} = \begin{bmatrix} d \\ e \end{bmatrix}$$

gdzie:

$$a = (B_x - A_x)^2 + (B_y - A_y)^2 + (B_z - A_z)^2$$

$$b = (B_x - A_x)(D_x - C_x) + (B_y - A_y)(D_y - C_y) + (B_z - A_z)(D_z - C_z)$$

$$c = -(D_x - C_x)^2 - (D_y - C_y)^2 - (D_z - C_z)^2$$

$$d = -(A_x - C_x)(B_x - A_x) - (A_y - C_y)(B_y - A_y) - (A_z - C_z)(B_z - A_z)$$

$$e = -(A_x - C_x)(D_x - C_x) - (A_y - C_y)(D_y - C_y) - (A_z - C_z)(D_z - C_z)$$

Obliczamy zatem wyznaczniki:

$$w = ac + b^2$$

$$w_m = cd + bd$$

$$w_n = ae - bd$$

Rozwiązanie tego układu jest jednoznaczne wtedy i tylko wtedy, gdy wiersze macierzy są liniowo niezależne, czyli gdy wyznacznik  $w$  jest różny od zera. Jest to równoważne z warunkiem, aby proste te nie były równoległe. Szukane parametry  $m$  i  $n$  określają punkty na obu prostych będące końcami najkrótszego łączącego je odcinka. Punkty te są równe:

$$M = A + m \cdot \vec{AB}$$

$$N = C + m \cdot \vec{CS}$$

Ostatnim krokiem do wyznaczenia punktu, w którym znajduje się obiekt, jest znalezienie środka powyższego odcinka:

$$O = \frac{M + N}{2} = \begin{bmatrix} \frac{M_x + N_x}{2} \\ \frac{M_y + N_y}{2} \\ \frac{M_z + N_z}{2} \end{bmatrix}$$

## **4 Analiza i przetwarzanie obrazu**

Aby obliczenia algebraiczne położenia obiektu w przestrzeni opisane w rozdziale 3. były możliwe, wcześniej musi zostać przeprowadzona analiza obrazu. Jej celem jest wyznaczenie punktów na obrazie, w których znajdują się środek obiektu oraz środek jego cienia. Zadanie to jest utrudnione przez jakość analizowanego zdjęcia. Podczas testów systemu zauważono, że najważniejszym powodem słabej jakości w przypadku wykorzystania sztucznego, punktowego źródła światła, jest niedostateczne doświetlenie sceny. Nie jest wówczas możliwe użycie kloszy lub wykorzystanie mocnego reflektora, co z pewnością poprawiłoby oświetlenie, gdyż pociągnęłoby to za sobą również utracenie pojedynczego, ostrego cienia. Próba uzyskania lepszej jasności zdjęcia poprzez wydłużenie czasu otwarcia migawki nie jest też możliwa, gdyż wykonywano zdjęcia obiektu ruchomego. Dłuższy czas powodowałby jego rozmazanie. W niniejszej pracy nie zostało znalezione rozwiązanie tego problemu i dlatego zdjęcia wykonywane przy świetle słonecznym pozwalają uzyskiwać dokładniejsze wyniki.

### **4.1 Algorytm przetwarzania obrazu**

Analiza obrazu okazuje się dawać najwięcej możliwości związanych z dostosowaniem systemu do warunków w jakich wykonywane są zdjęcia. To właśnie ona ma największy wpływ na dokładność całego system.

Całość użytego podczas testów algorytmu przedstawiono schematycznie na Rys. 13. Pierwszym jego elementem jest wczytanie obrazu. Napisana aplikacja umożliwia jedynie wczytanie zdjęć z plików, lecz istnieje możliwość implementacji i dodania do niej obsługi kamery. Do dalszej obróbki, obraz powinien zostać przekazany w postaci monochromatycznej. Może być to uzyskane poprzez rejestrowanie obrazu czarno-białego przez kamerę lub cyfrowe usunięcie koloru. Kolejnym krokiem jest zastosowanie maski. Pozwala to na zmniejszenie obszaru analizowanego zdjęcia do części, w której może pojawić się szukany obiekt. Pozwala to zarówno skrócić czas obliczeń, jak i wyeliminować z obrazu elementy

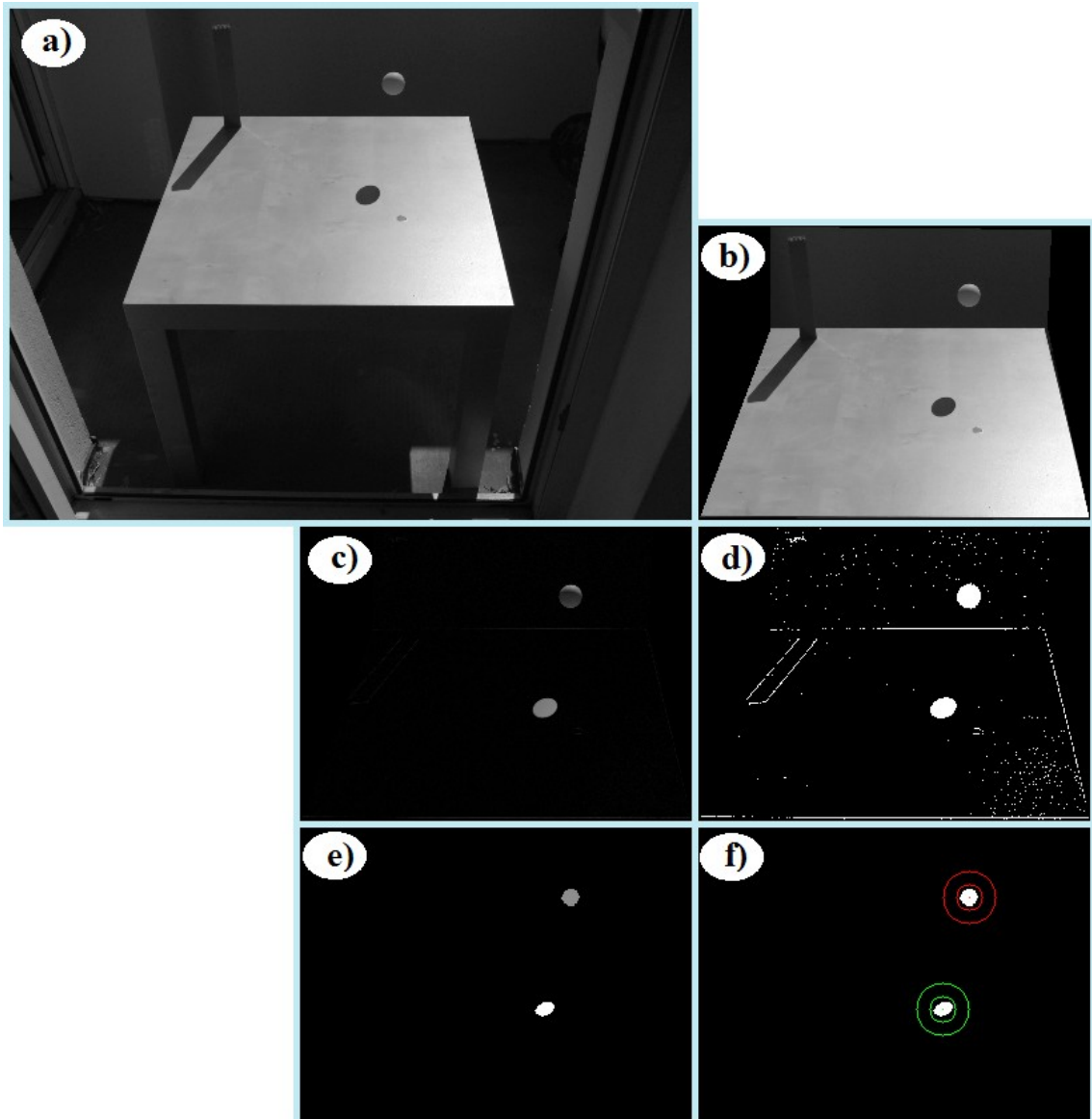


zakłócające dalszą jego obróbkę.



Rys. 13: Schemat zastosowanego algorytmu

Następnie wykonywane jest odejmowanie zapisanego wcześniej tła od obrazu i binaryzacja wyniku. Po tej operacji na obrazie zostaje jedynie obiekt, cień i niewielkie szumy, które usuwane są przy użyciu erozji. Ostatnim krokiem jest wyznaczenie środków obu obiektów na obrazie. Ważne jest, aby poprawnie określić, który ze znalezionych punktów jest środkiem cienia, a który obiektu. Przyjęte zostało założenie, że obiekt jest zawsze wyżej niż cień. W ogólnym przypadku nie musi to być prawda. Dla każdej sceny, położenia aparatu i źródła światła można jednak określić zasadę wzajemnego położenia cienia i obiektu, nawet jeżeli jest konieczność obliczania tego na podstawie czasu (przy świetle słonecznym) czy położenia na obrazie. Rys. 14 przedstawia jeden obraz po kolejnych opisanych przekształceniach.



Rys. 14: Zmiany obrazu po kolejnych krokach algorytmu przetwarzania:  
a) usunięcie koloru, b) zastosowanie maski, zmniejszenie obszaru c) odjęcie tła,  
d) binaryzacja, e) erozja i oznaczenie obiektów, f) wyznaczenie środków obiektów

## 4.2 Opis poszczególnych kroków przetwarzania

**Usunięcie koloru** konwertuje zarejestrowany obraz kolorowy na czarno-biały. Mimo, że zmniejsza to ilość informacji zapisanych w obrazie, co w przypadku pewnych scen mogłoby uniemożliwić poprawną analizę, to właśnie obrazy

monochromatyczne są łatwiejsze w obróbce i trwa ona krócej. Operacja usuwania koloru podczas przeprowadzanych testów była przeprowadzana na etapie wczytywania zdjęć z nośnika danych. Może być ona również wykonywana jako osobny krok lub użyta może zostać kamera, która rejestruje zdjęcia w odcieniach szarości. Takie rozwiązanie mogłoby obniżyć cenę systemu wizyjnego (zakup kamery) oraz zmniejszyć ilość przesyłanych lub zapisywanych danych.

Kamery i aparaty cyfrowe zazwyczaj zapisują zdjęcia kolorowe w przestrzeni barw RGB co jest naturalnym rozwiązaniem wynikającym z wewnętrznej budowy matrycy CCD stosowanej w tych urządzeniach. Konwersja obrazu zapisanego w przestrzeni barw RGB na obraz czarno-biały wykonywana jest przy użyciu biblioteki OpenCV [4] jako przekształcenie punktowe opisane wzorem:

$$Y = 0,299 R + 0,578 G + 0,114 B$$

gdzie:

$Y$  – jasność piksela na obrazie monochromatycznym

$R$  – składowa *czzerwona* piksela obrazu kolorowego

$G$  – składowa *zielona* piksela obrazu kolorowego

$B$  – składowa *niebieska* piksela obrazu kolorowego

**Zastosowanie maski** pozwala na wybranie z obrazu tylko tej jego części, na której znajdować się mogą obiekty interesujące z punktu widzenia działania systemu. Pozwala to uniknąć zakłóceń na obrazie, wynikających z pojawienia się w kadrze dodatkowych elementów, np. inne ruszające się obiekty, przechodzący ludzie itp. Zazwyczaj stosowana maska przy obróbce obrazu jedynie przysłania wybrane fragmenty obrazu powodując, że są one analizowane lecz nie przynoszą żadnych wyników. Moduł stosujący maskę zastosowany w używanym algorytmie przycina również obraz tak, aby nie było na nim usuniętych przez maskę całych wierszy lub kolumn. Zmniejsza to rozmiar obrazu pozwalając na skrócenie czasu obliczeń.

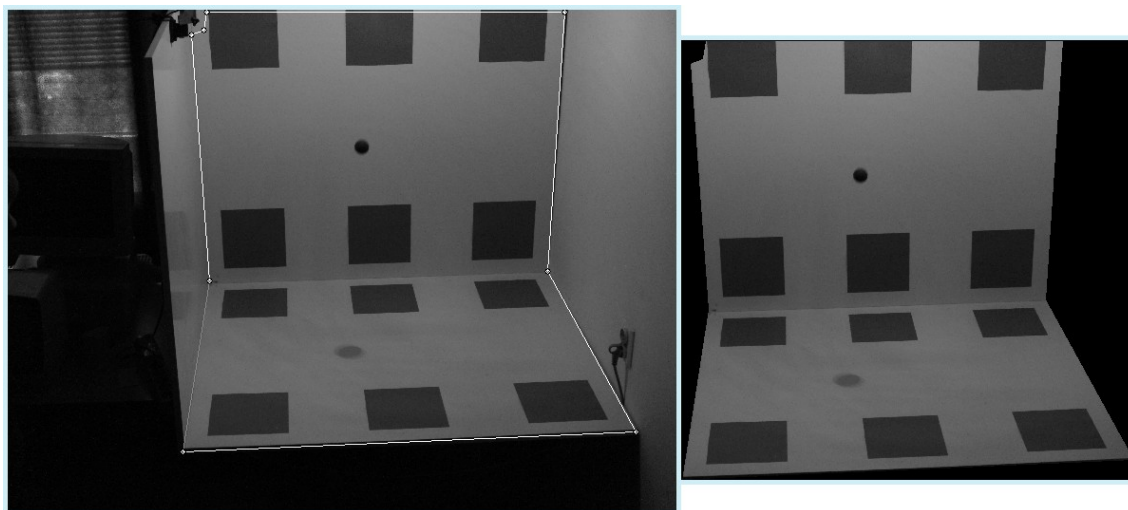
---

## Wykorzystanie cienia do odtworzenia pozycji obiektu w przestrzeni

---

Maska jest obrazem binarnym, gdzie piksele o wartości 0 oznaczać będą piksele do usunięcia, a piksele o wartości 1 - piksele, które mają pozostać na obrazie. Przyjęcie takiej formy zapisu maski pozwala na zastosowanie jej poprzez pomnożenie odpowiadających sobie pikseli na obu obrazach. Wówczas część obrazu pozostaje niezmienną, a druga, niepotrzebna część, usunięta – pikselom przypisana jest wartość 0. Inną metodą, niewymagającą wykonywania działania mnożenia, jest przypisanie pikselom wartości zerowej lub pozostawienie istniejącej, w zależności od stanu odpowiadających im pikseli na obrazie maski.

Rys. 15 przedstawia oryginalny obraz z naniesionymi punktami ograniczającymi kształt maski oraz obraz po jej zastosowaniu. Analizowany obszar został pomniejszony do płaszczyzny podstawy oraz płaszczyzny pionowej, a przysłonięte zasłoną okno, znajdujące się po prawej stronie zdjęcia – usunięte. Mogłoby ono powodować zakłócenia wynikające ze zmiany jasności, a dodatkowo niepotrzebnie wydłużałoby czas obliczeń – obiekt nigdy nie mógłby znaleźć się na jego tle ze względu na ustawienie sceny.



Rys. 15: Analizowany obraz przed i po zastosowaniu maski usuwającej zbędne obszary

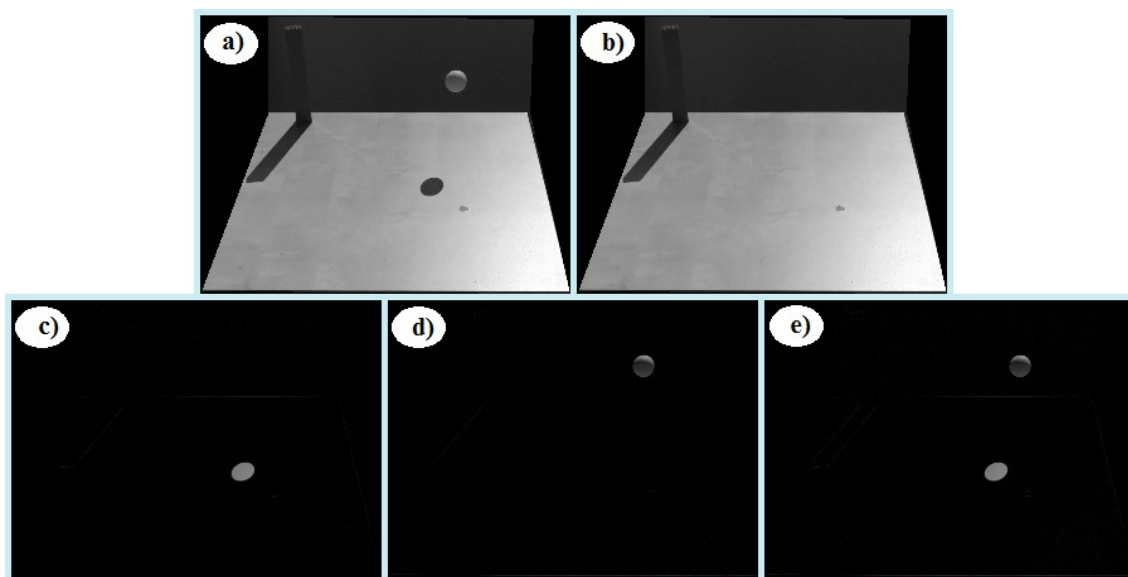
**Odjęcie tła** wydaje się być najprostszą metodą odnalezienia na obrazie

szukanych obiektów. Konieczne jest wykonanie i zarejestrowanie zdjęcia tła przed wejściem w przestrzeń sceny szukanych obiektów. Następnie, tło odejmowane jest od przetwarzanego obrazu. W założeniu, po tej operacji na obrazie pozostają jedynie poszukiwane obiekty lecz w rzeczywistości, widoczne są również naturalnie powstające szумы.

Ważnym aspektem wartym rozważenia jest sposób odejmowania. Problem związany jest z reprezentacją liczbową obrazu – są to 8-bitowe zmienne liczbowe bez znaku, czyli zakres ich wartości wynosi  $0 \div 255$ . Gdy wynik odejmowania przekroczy dolny zakres, pikselowi przypisywana jest wartość 0. Może to skutkować utratą informacji zawartą na obrazie – „zniknięciem” całego obiektu. W związku z powyższym, pod uwagę brane są trzy sposoby odejmowania:

- odejmowanie tła od obrazu,
- odejmowanie obrazu od tła,
- wartość bezwzględna odejmowania obu obrazów.

Na Rys. 16 przedstawione są dwa zarejestrowane zdjęcia: analizowany obraz i wcześniej zarejestrowane tło, oraz wyniki ich odejmowania na trzy wymienione sposoby. Należy zauważyć, że na Rys. 16 c) został wykryty jedynie cień, a na obrazie Rys. 16 d) jedynie obiekt. Jest to spowodowane wspomnianym wcześniej przekroczeniem dolnego zakresu wartości zmiennej przez otrzymany wynik odejmowania. Problem wydaje się rozwiązywać obliczanie wartości bezwzględnej odejmowania, gdyż widoczne są wówczas oba elementy, co pokazuje Rys. 16 e). Pojawia się jednak większa ilość zakłóceń, będąca sumą zakłóceń na obrazach uzyskanych dwoma wcześniejszymi metodami odejmowania.

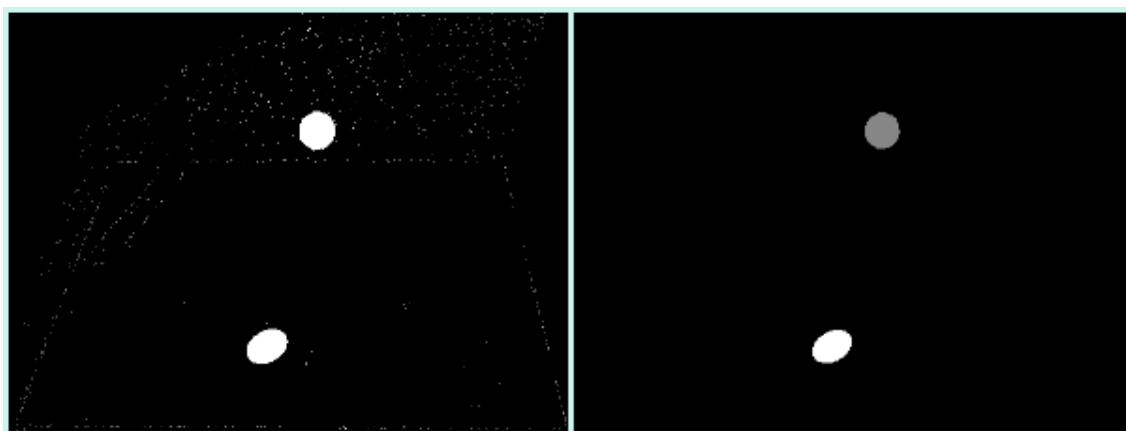


Rys. 16: Odejmowanie tła: a) analizowany obraz, b) obraz tła, c) odejmowanie tła od obrazu, d) odejmowanie obrazu od tła, e) wartość bezwzględna odejmowania

**Binaryzacja** jest jednym z najprostszych przekształceń punktowych, a zarazem jest ona bardzo często używana podczas przetwarzania obrazu. Większość pomiarów i analiz przeprowadzana jest właśnie na obrazie binarnym i często możliwa jest tylko na nim [5].

Podstawowym problemem, pojawiającym się przy wykonywaniu tej operacji, jest znalezienie progu binaryzacji. W celu jego wyznaczenia często wykorzystywany jest histogram obrazu. Tło sceny dobierane jest zazwyczaj w taki sposób, aby kontrastowało z obiektem, co pozwala na odczytanie z histogramu wartości, która rozdziela piksele tła od obiektu. Gdy odejmowane jest od analizowanego obrazu, jak ma to miejsce w prezentowanym algorytmie, na obrazie powinny na nim zostać już jedynie szukane obiekty. Binaryzacja wykonywana jest więc jedynie do usunięcia powstałych szumów i innych zakłóceń, więc wartość progu odcinania jest bardzo niska. Dodatkowo, testy pokazały, że nie warto dążyć do usunięcia wszystkich szumów przy użyciu binaryzacji ponieważ powoduje to zniekształcenie obiektów. Lepszym rozwiązaniem jest pozostawienie pewnej ilości szumów i usunięcie ich przy pomocy erozji.

**Erozja** jest podstawowym przekształceniem morfologicznym, które może być interpretowane jako filtr minimalny – każdemu pikselowi przypisywane jest minimum z wartości pikseli znajdujących się w jego otoczeniu. Erozja wykonywana na obrazie binarnym pozwala na usunięcie powstałych podczas wcześniejszych przekształceń szumów oraz innych, mniejszych artefaktów. Należy zauważyć, że przekształcenie to zniekształca również poszukiwane na obrazie obiekty, jednak przy założeniu, że szum jest zawsze znacznie mniejszy, zniekształcenie to jest niewielkie. Istotną kwestią jest tutaj wybór elementu strukturalnego, który determinuje otoczenie punktu. Podczas testów rozważane były elementy kwadratowe o boku 3, 5 i 7 pikseli, oraz „koła” wpisane w te kwadraty. Zgodnie z teorią, można zauważyć addytywność tego przekształcenia – powiększenie rozmiaru elementu strukturalnego pozwala osiągnąć ten sam lub bardzo podobny efekt po mniejszej ilości następujących po sobie erozji. Kształt elementu został wyznaczony eksperymentalnie i wybrany został kwadrat, który w przypadku analizowanych zdjęć okazał się powodować ostatecznie mniejsze zniekształcenia obiektów niż element w kształcie „koła”. W stworzonej aplikacji erozja wykonywana jest kilkakrotnie na przemian z opisanym poniżej etykietowaniem, aż do osiągnięcia dwóch elementów na obrazie, czyli cienia i obiektu, lub do czasu przekroczenia maksymalnej dozwolonej liczby powtórzeń.



*Rys. 17: Obraz binarny i ten sam obraz po wykonaniu erozji i indeksowaniu obiektów*

**Indeksowanie obiektów** jest operacją, która pozwala na oznaczenie na obrazie poszczególnych obiektów, rozróżniając je od siebie poprzez przypisanie należącym do nich pikselom charakterystycznych wartości. Drugą informacją pozyskiwaną podczas wykonywania tego przekształcenia jest wyznaczenie liczby obiektów znajdujących się na obrazie.

Indeksowanie wykonywane jest poprzez analizę kolejnych pikseli obrazu i ich otoczenia. Początkowo, punkty na obrazie posiadają dwie możliwe wartości, wskazujące czy należą one do tła, czy do jednego z obiektów. Następnie, analizując obraz punkt po punkcie, pikselom obiektów przypisywane są nowe wartości czyli indeksy. Nadawane są one w zależności od ich otoczenia:

- jeżeli w otoczeniu nie znajdują się jeszcze indeksowane piksele, wybierana jest wartość o jeden większa od maksymalnego użytego do tej pory indeksu,
- jeżeli w otoczeniu znajduje się piksel o nadanym indeksie, wybierana wartość jest powieleniem tego indeksu.

Taki algorytm powoduje jednak, że w bezpośrednim sąsiedztwie mogą znaleźć się piksele o różnych indeksach. Kolejnym krokiem musi być zatem wyeliminowanie tego zjawiska. W tym celu używana jest tablica sklejeń. Jest to tablica zawierająca informacje o różnych indeksach nadanych pikselom tego samego obiektu. Pozwala ona na modyfikację nadanych indeksów w taki sposób, aby występowało jednoznaczne przyporządkowanie pomiędzy indeksem a obiektem.

Liczba obiektów występujących na obrazie może zostać wyznaczona, przy zachowaniu odpowiedniego algorytmu, jako najwyższy nadany indeks podczas wykonywania etapu sklejanego.

**Wyznaczenie środków znalezionych obiektów** jest końcowym etapem analizy obrazu. Po rozróżnieniu obiektów, poprzez nadanie ich pikselom odpowiednich wartości, należy zredukować ilość informacji o nich, do dwóch współrzędnych definiujących jednoznacznie położenie całego obiektu na obrazie.



Pierwszym nasuwającym się sposobem wyznaczenia środka obiektu, jest znalezienie jego środka ciężkości, czyli momentów bezwładności pierwszego rzędu [5]:

$$M_{1x} = \frac{1}{A(F)} \sum_F x_i$$

$$M_{1y} = \frac{1}{A(F)} \sum_F y_i$$

gdzie:

$F$  - analizowana figura

$A(F)$  - pole powierzchni figury (liczba punktów figury)

$x_i, y_i$  - współrzędne punktów figury

Metoda ta jest stosunkowo dokładna w przypadku, gdy obiekt został wyraźnie odwzorowany na obrazie. Ze względu na prostotę i przejrzystość to ona była stosowana podczas testów.

### **4.3 Inne koncepcje przetwarzania obrazu**

Użyty algorytm przetwarzania obrazu wykorzystuje proste i podstawowe metody w celu wyznaczenia pozycji szukanych obiektów. Pozwala to stworzyć poprawnie działający system wyznaczający pozycję bez zagłębiania się w skomplikowane zagadnienia. Aby zwiększyć dokładność obliczeń, a przede wszystkim, aby można było stosować system w trudniejszych warunkach oświetleniowych, należy rozwinąć go o bardziej zaawansowane metody.

Jednym ze źródeł niedokładności wyznaczania pozycji jest błędne znalezienie środka obiektu na obrazie. Poprawienie tego kroku może zostać uzyskane poprzez wykorzystanie informacji o kształcie obiektu. W przypadku piłki używanej podczas testów dobrym rozwiązaniem mogłoby być przybliżanie odnalezionego na obrazie obiektu elipsą, a następnie znalezienie jej środka. Postępowanie takie pozwoliłoby zmniejszyć negatywny wpływ odbłasków i innych

niewielkich artefaktów pojawiających się na obrazie. Przybliżanie elipsą mogłoby także wyeliminować konieczność przeprowadzania erozji. Należałoby w takim przypadku wyznaczyć elipsy dla wszystkich obiektów, a następnie wybrać spośród nich dwa o dostatecznie dużym promieniu.

Najtrudniejszym elementem podczas analizy obrazu wydaje się być jednak wydzielenie obiektów na obrazie. Szczególnie dużym utrudnieniem jest zmieniające się, nawet w niewielkim stopniu, oświetlenie. Rozwiązaniem może być wykorzystanie algorytmów generacji tła, które pozwalają na dynamiczne tworzenie obrazu tła [6]. Uwzględniać mogą one zmieniające się oświetlenie, pozostawione przedmioty, czy pojawiający się brud. Dzięki temu, artefakty te nie są oznaczane jako szukane obiekty.

Innym, prostszym rozwiązaniem, może być próba wyodrębnienia obiektów na obrazie wykorzystując wartości koloru. W tym celu należałoby zmienić przestrzeń barw na HSV i wykorzystując składową barwy, wykonać segmentację. Utrudnieniem w tej metodzie okazuje się być cień. Teoretycznie nie zmienia on barwy podłoża, na które pada, a jedynie nasycenie. Należałoby zatem wykorzystać inny algorytm do znalezienia cienia, a inny do znalezienia obiektu na obrazie.

## **5 Opis aplikacji**

### **5.1 Postawione wymagania**

Głównym celem stworzonej aplikacji była możliwość przeprowadzenia badań na temat możliwości wykorzystania cienia do odtworzenia pozycji w przestrzeni oraz ocenę jakości takiego rozwiązania. Nie musi więc ona spełniać wymagań, które postawione byłyby przed zadaniem tworzenia systemu wizyjnego, mającym działać w produkcji. Nie jest to więc system czasu rzeczywistego, który miałby narzucone wymogi czasowe, pewność i przewidywalność działania. Obraz nie jest także przetwarzany w czasie rzeczywistym – analizowane są jedynie wcześniej przygotowane serie zdjęć. Podstawową cechą ma być natomiast łatwość zmiany procesu przetwarzania obrazu, aby umożliwić przetestowanie różnych konfiguracji.

Wymagania postawione wobec aplikacji:

1. Obliczanie pozycji obiektu w przestrzeni na podstawie cienia (przy świetle słonecznym oraz sztucznym – żarówka).
2. Łatwość zmiany procesu przetwarzania obrazu oraz jego testowania:
  - a) edytowalna lista procesów,
  - b) możliwość łatwej zmiany parametrów procesów przetwarzania,
  - c) podgląd przetwarzanego obrazu po każdym wykonanym kroku.
3. Zapisywanie i otwieranie zapisanych konfiguracji.
4. Podgląd i eksport wyników.
5. Modułowość – możliwość dodania innych sposobów analizy obrazu.

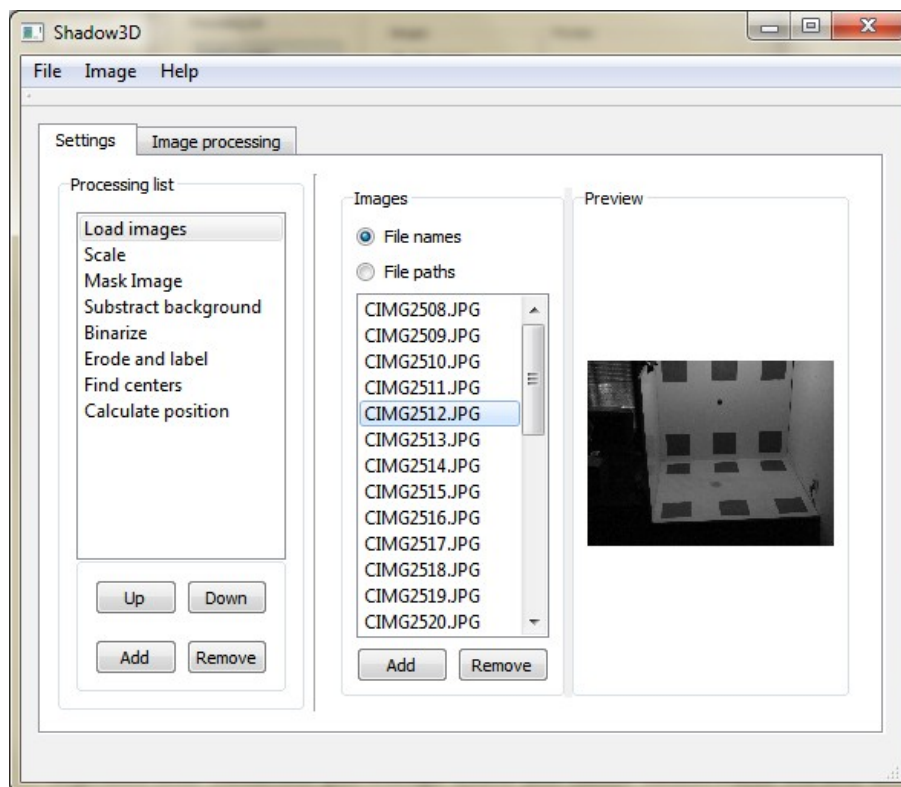
### **5.2 Ogólny opis działania**

Aplikacja została podzielona na graficzny interfejs użytkownika oraz na część zasadniczą, analizującą obraz i obliczającą pozycję. Cała aplikacja jest

napisana w języku C++ [7] z wykorzystaniem biblioteki QT [8]. Wykorzystano wiele jej elementów składowych. Przede wszystkim pozwoliła ona na stworzenie graficznego interfejsu użytkownika (GUI) oraz na oddzielenie wątku obliczeniowego od wątku służącego do obsługi GUI [9]. Funkcje zapisu danych do pliku umożliwiły przechowywanie używanych konfiguracji oraz eksportu danych. Graficzny interfejs użytkownika ułatwia zmianę parametrów poszczególnych procesów przetwarzania obrazu. Pozwala on również na szybki podgląd kolejnych kroków analizy. Przetwarzanie obrazów wykonywane jest z wykorzystaniem biblioteki Open CV [10].

Podstawowym zadaniem aplikacji jest obliczenie pozycji obiektu w przestrzeni. W tym celu konieczne jest wyznaczenie położenia szukanego elementu i jego cienia na obrazie oraz wykonanie obliczeń geometrycznych. Aby proces przetwarzania obrazu był łatwo edytowalny, poszczególne kroki zostają zapisane w postaci listy kolejnych modułów analizujących obraz. Pozwala to na sprawdzenie różnych konfiguracji systemu oraz tworzenie nowych modułów przetwarzania obrazu. Dowolność taka niesie jednak za sobą pewne niebezpieczeństwo. Wymagana jest znajomość aplikacji przez użytkownika, aby działanie wprowadzonego algorytmu miało sens. Przykładem może być konieczność rozpoczęcia listy procesów przetwarzania obrazu od modułu wczytującego obraz i zakończenie modułem obliczającym położenie w przestrzeni. Lista kolejnych procesów analizujących obraz jest połączona z listą paneli służących do ich obsługi poprzez panel interfejsu graficznego. Jest to przyporządkowanie jeden do jeden – każdy proces posiada osobny, odpowiadający mu panel, pozwalający na zmianę jego ustawień.

Na Rys. 18 przedstawiony został widok aplikacji. Po lewej stronie widoczna jest edytowalna lista procesów przetwarzania. Wybranie jednego z nich powoduje wyświetlenie ustawień i parametrów w odpowiadającym mu panelu. Dodatkowo, zazwyczaj wyświetlany jest również podgląd obrazu, aby możliwa była analiza i ocena poszczególnych kroków przetwarzania.



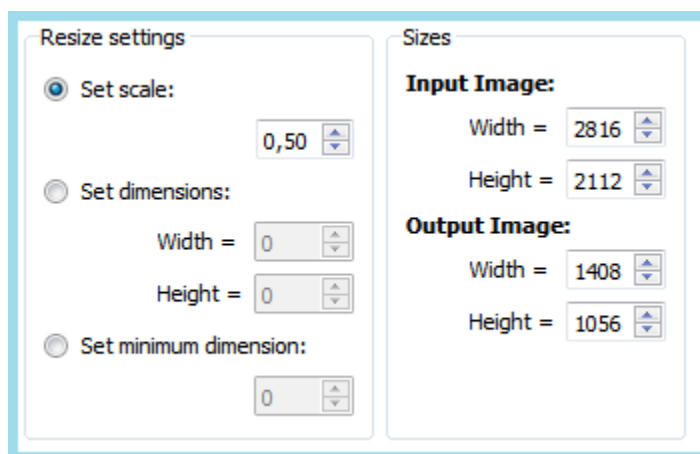
Rys. 18: Widok ogólny aplikacji – lista kroków przetwarzania obrazu oraz panel wczytywania obrazów

### 5.3 Ustawienia poszczególnych procesów analizy obrazu

Lista kolejno następujących po sobie procesów przetwarzania tworzy opis całej analizy obrazu – od jego wczytania, aż po uzyskanie położenia obiektu w przestrzeni. Zadaniem pierwszego z nich jest zatem załadowanie obrazu, który następnie będzie przekazywany do kolejnych procesów. Zaprezentowany na Rys. 18 panel ustawień pozwala na wybór poszczególnych zdjęć z dysku lub innego nośnika oraz na ich podgląd. Zaznaczenie konkretnego zdjęcia na tym panelu, powoduje, że to ono jest analizowane podczas kolejnych kroków przetwarzania. Dodatkowo, skróty klawiszowe [Ctrl+↓] i [Ctrl+↑] pozwalają na wybór odpowiednio następnego i poprzedniego zdjęcia z tej listy. Dostępne są one z każdego miejsca w programie, czyli również z paneli innych procesów przetwarzania. Ułatwia to analizę wprowadzonych parametrów z użyciem różnych obrazów. Moduł ten mógłby być również zastąpiony innym, który pobiera zdjęcia

z kamery w czasie rzeczywistym.

Następnym zaimplementowanym modułem jest skalowanie zdjęcia (jego widok został przedstawiony na Rys. 19). Pozwala on na zmniejszenie analizowanego obrazu, co znacząco przekłada się na czas obliczeń. Stworzone zostały trzy sposoby skalowania obrazu. Pierwszy z nich to skalowanie obu wymiarów przez ten sam czynnik. Drugi jest przeskalowaniem obrazu osobno w dwóch wymiarach tak, aby rozmiar był równy podanemu. Trzeci sposób pozwala ustalić mniejszy z wymiarów obrazu, podczas gdy drugi wymiar zostanie obliczony w sposób zachowujący skalę. Sposoby te w pewnym stopniu zawierają powielenie swojej funkcjonalności – np. wszystkie operacje można wykonać z użyciem drugiej opcji. Pozwalają jednak użytkownikowi na łatwiejsze wprowadzenie parametrów, w zależności od jego preferencji.

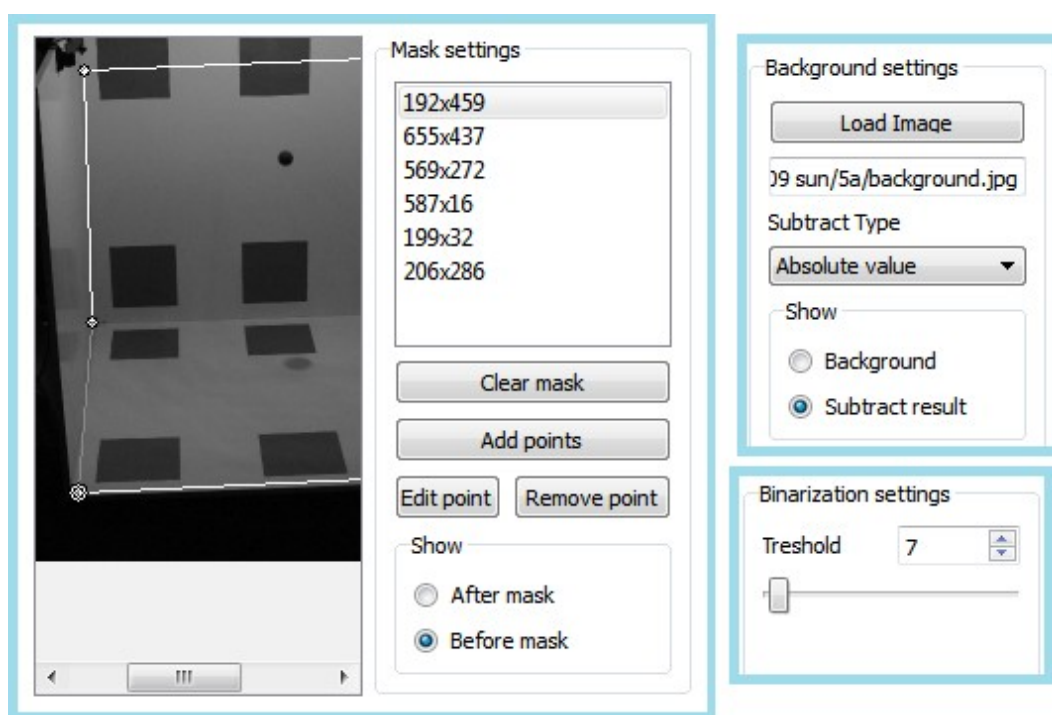


Rys. 19: Widok ustawień procesu skalowania

Kolejne trzy panele zostały przedstawione na Rys. 20. Na pierwszym z nich widoczne są ustawienia nakładanej maski. Definiowana jest ona przez listę kolejnych punktów tworzących wielobok. Wnętrze tej figury uznawane jest za obszar, w którym może znajdować się poszukiwany obiekt i tylko on jest przetwarzany. Punkty te wskazywane są na pełnym obrazie, ale dostępny jest również podgląd obrazu po nałożeniu maski i przycięciu jego rozmiaru. Możliwe

jest również usunięcie całej maski oraz edycja poszczególnych punktów.

Drugim panelem przedstawionym na Rys. 20 są ustawienia procesu odejmującego tło. Poza wyborem obrazu przedstawiającego tło, które musi zostać wczytane z dysku, konieczny jest również wybór trybu odejmowania – odejmowanie tła od obrazu, obrazu od tła oraz wartość bezwzględna odejmowania. Różnice pomiędzy tymi opcjami opisane zostały w podrozdziale 4.2.

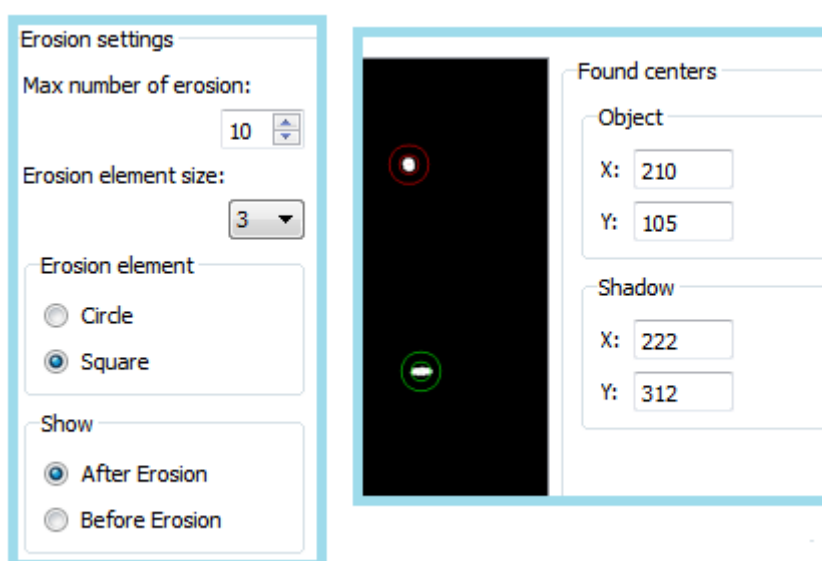


Rys. 20: Widok paneli ustawień: nakładania maski, odejmowania tła i binaryzacji

Na Rys. 20 przedstawiono również ustawienia procesu binaryzacji. W rzeczywistości możliwa do zmiany jest tylko jedna wartość – próg binaryzacji. Tak mało ustawień jest związane z prostotą tego procesu. Rozważana była również binaryzacja odwrotna oraz przedziałowa, jednak nie zostały one zaimplementowane ponieważ nie znaleziono dla nich praktycznego zastosowania w tej wersji aplikacji.

Panele dwóch ostatnich procesów przetwarzających obraz zostały

przedstawione na Rys. 21. Pierwszy z nich, czyli erozja, połączona jest indeksowaniem elementów tak, aby dążyć do pozostawienia dokładnie dwóch elementów na obrazie wynikowym. Parametry możliwe do zmiany dotyczą jedynie erozji. Jest to kształt elementu erodującego – kwadrat lub koło oraz jego rozmiar. Dodatkowo występuje ograniczenie maksymalnej ilości iteracji wykonywania tego procesu w przypadku pozostawienia większej ilości elementów na obrazie.



Rys. 21: Panele ustawień: erozja i indeksowanie oraz wyznaczanie środków obiektów

Poszukiwanie środków obiektów nie posiada żadnych parametrów do zmiany. Panel dotyczący tego procesu służy jedynie celom informacyjnym. Wyświetlenie obliczonych wartości, wraz z oznaczeniem punktów na obrazie kolorowymi okręgami, pozwala stwierdzić poprawność działania całego algorytmu przetwarzania obrazu.

#### **5.4 Obliczenia i prezentacja wyników**

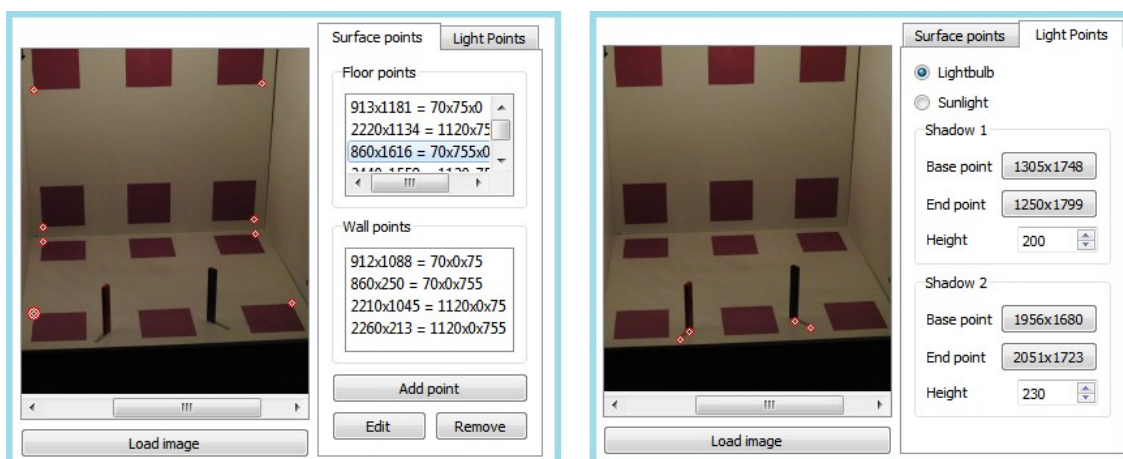
Ostatni krok algorytmu nie jest w rzeczywistości procesem przetwarzającym obraz. Przeprowadzane są w nim jedynie obliczenia korzystające



## Wykorzystanie cienia do odtworzenia pozycji obiektu w przestrzeni

z uzyskanych wcześniej z obrazu danych – położenia obiektu i cienia na obrazie oraz przeprowadzonej kalibracji. Do obliczeń wykorzystana została oddzielna biblioteka posiadająca rozbudowane narzędzia do algebry liniowej – Armadillo [11].

Panel ustawień tego procesu pozwala zatem na kalibrację systemu i został przedstawiony na Rys. 22. Do wyznaczenia współczynników transformacji płaskich należy wprowadzić dwie listy punktów, osobno dla podstawy i osobno dla płaszczyzny pionowej. Każdy punkt niesie za sobą dwie informacje – jego położenie na obrazie wybierane na podglądzie poprzez kliknięcie przez użytkownika oraz jego rzeczywiste współrzędne w przestrzeni wpisywane w oknie dialogowym. Wymagane jest podanie czterech punktów dla każdej płaszczyzny, z których żadne trzy nie są współliniowe. Istnieje również możliwość wprowadzenia większej ilości punktów, co poprawia dokładność wykonanej kalibracji.



Rys. 22: Panel kalibracji systemu

Druga część kalibracji, również pokazana na Rys. 22, pozwala na określenie położenia światła. Polega ona na wskazaniu cienia rzucanego przez obiekt, a dokładniej jego początku oraz końca (Rys. 23). Istotna jest również znajomość

---

## Wykorzystanie cienia do odtworzenia pozycji obiektu w przestrzeni

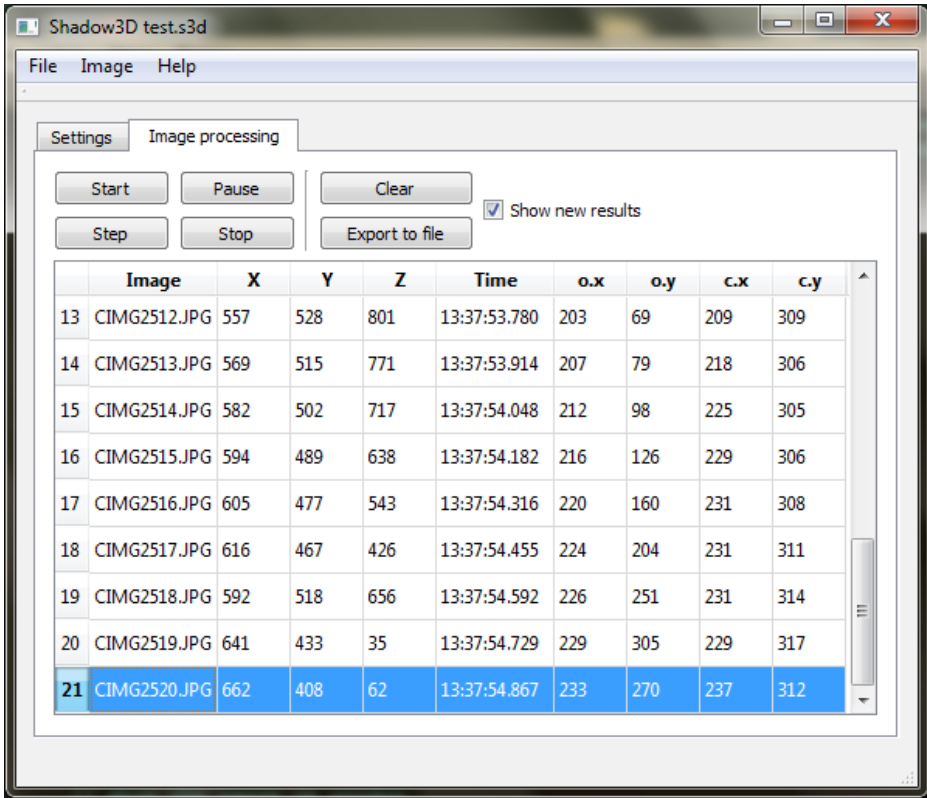
---

wysokości tego obiektu. Ważnym parametrem jest wybór natury światła – punktowe światło żarówki lub światło słoneczne. Przy wyborze światła punktowego wymagane są dwa obiekty kalibrujące, rzucające dwa cienie, które nie są równoległe względem siebie. Przy świetle słonecznym wystarczający jest jeden cień. Jest to związane z różnym rozchodzeniem się promieni i powiązaną z tym różną geometrią analizowanej sceny.



Rys. 23: Punkty kalibracji cienia

Ostatni panel użytkownika, przedstawiony na Rys. 24, daje możliwość uruchomienia przetwarzania obrazu oraz pozwala na wyświetlanie wyników. Przycisk *Start* rozpoczyna analizę kolejnych obrazów, które są dostarczane przez moduł wczytujący. Może on zatem zakończyć się po ostatnim wybranym zdjęciu lub trwać nieprzerwanie, w przypadku gdyby został stworzony i wybrany moduł wczytujący dane z kamery. Przycisk *Step* pozwala na krokowe przetwarzanie – każde kliknięcie spowoduje analizę tylko jednego, kolejnego obrazu. Przetwarzanie serii można również zatrzymać przyciskami *Pause* oraz *Stop*.



The screenshot shows the 'Shadow3D test.s3d' application window. It has a menu bar with 'File', 'Image', and 'Help'. Below the menu bar are two tabs: 'Settings' and 'Image processing'. The 'Image processing' tab is active and contains several buttons: 'Start', 'Pause', 'Clear', 'Step', 'Stop', and 'Export to file'. There is also a checkbox labeled 'Show new results' which is checked. Below the buttons is a table with 10 columns: 'Image', 'X', 'Y', 'Z', 'Time', 'o.x', 'o.y', 'c.x', and 'c.y'. The table contains 10 rows of data, with the last row (row 21) highlighted in blue.

	Image	X	Y	Z	Time	o.x	o.y	c.x	c.y
13	CIMG2512.JPG	557	528	801	13:37:53.780	203	69	209	309
14	CIMG2513.JPG	569	515	771	13:37:53.914	207	79	218	306
15	CIMG2514.JPG	582	502	717	13:37:54.048	212	98	225	305
16	CIMG2515.JPG	594	489	638	13:37:54.182	216	126	229	306
17	CIMG2516.JPG	605	477	543	13:37:54.316	220	160	231	308
18	CIMG2517.JPG	616	467	426	13:37:54.455	224	204	231	311
19	CIMG2518.JPG	592	518	656	13:37:54.592	226	251	231	314
20	CIMG2519.JPG	641	433	35	13:37:54.729	229	305	229	317
21	CIMG2520.JPG	662	408	62	13:37:54.867	233	270	237	312

Rys. 24: Widok panelu obliczeń i prezentacji wyników

Kolejny przycisk – *Clear* pozwala na wyczyszczenie listy zapamiętanych wyników, natomiast przycisk *Export* pozwala na zapis wyników przetwarzania do pliku. Dane te zapisywane są w postaci tekstowej – jako liczby. Dzięki temu możliwe jest wczytanie ich do innych programów w celu dalszej analizy, np. Excel lub Matlab [12], który był wykorzystywany podczas przeprowadzanych testów.

Format zapisu ma umożliwić automatyczne wczytanie danych przez inne aplikacje. Po pierwszej linijce nagłówka znajdują się kolejne rekordy danych oddzielone od siebie znakiem końca linii. Dane liczbowe w każdym rekordzie oddzielone są od siebie przecinkami. Wszystkie zapisywane wartości są liczbami całkowitymi. W każdym rekordzie znajdują się kolejno następujące dane:

- numer przetwarzanego obrazu,
- współrzędna  $X$  położenia obiektu w przestrzeni [mm],

## Wykorzystanie cienia do odtworzenia pozycji obiektu w przestrzeni

---

- współrzędna  $Y$  położenia obiektu w przestrzeni [mm],
- współrzędna  $Z$  położenia obiektu w przestrzeni [mm],
- pozycja obiektu na obrazie – współrzędna  $X$ ,
- pozycja obiektu na obrazie – współrzędna  $Y$ ,
- pozycja cienia na obrazie – współrzędna  $X$ ,
- pozycja cienia na obrazie – współrzędna  $Y$ ,
- czas wczytania zdjęcia – godzina,
- czas wczytania zdjęcia – minuty,
- czas wczytania zdjęcia – sekundy,
- czas wczytania zdjęcia – milisekundy.

Gdy analiza obrazu nie powiedzie się, wystąpi błąd podczas przetwarzania obrazu, jej wynik nie będzie zapisywany. Taki błąd można zauważyć jedynie poprzez brak rekordu o kolejnym numerze. Poniżej przedstawiono przykładowy plik zapisu. Jest to plik z danymi analizowanymi w jednym z testów, a wykresy z niego wygenerowane przedstawione zostały na Rys. 26 (str. 57). Można zauważyć brak danych dla obrazu nr 7 oraz 19. Odpowiadające im punkty nie zostały również zaznaczone na wspomnianym wykresie.

No	X	Y	Z	ox	oy	sx	sy	hh	mm	ss	zzzz
0	34,416	32,26	68,28	79,12	01,49	589					
1	58,391	104,30	46,47	80,12	01,50	018					
2	80,386	153,37	28,62	79,12	01,50	456					
3	102,384	177,44	19,75	78,12	01,50	895					
4	123,384	178,52	18,84	77,12	01,51	336					
5	147,381	153,62	27,90	77,12	01,51	762					
6	170,375	102,72	45,91	78,12	01,52	183					
8	216,363	93,90	48,107	78,12	01,52	672					
9	237,359	135,97	33,123	77,12	01,53	105					
10	259,357	150,105	27,135	76,12	01,53	540					
11	280,353	142,113	30,142	76,12	01,53	972					

---

Wykorzystanie cienia do odtworzenia pozycji obiektu w przestrzeni

---

12,302,344,112,122,41,145,77,12,01,54,402
13,334,338,35,135,67,142,78,12,01,54,833
14,355,337,54,142,60,154,77,12,01,55,265
15,370,333,111,148,40,171,76,12,01,55,699
16,391,324,129,156,34,184,76,12,01,56,134
17,413,319,121,164,37,191,76,12,01,56,559
18,440,317,82,174,50,192,76,12,01,56,982
20,487,307,59,191,58,206,76,12,01,57,483

## 6 Testy systemu wizyjnego i ich wyniki

Wszystkie testy przeprowadzono w oparciu o wykonane przez autora niniejszej pracy zdjęcia. Wykorzystany został aparat CASIO EX-F1 [13]. Wybrane dane techniczne zostały zawarte w Tabeli 1:

*Tabela 1: Wybrane dane techniczne użytego aparatu CASIO EX-F1*

Model:	CASIO EX-F1
Format zapisu zdjęć	RAW (DNG), JPEG, DCF 1.0 standard
Efektywne piksele	6.0 megapikseli
Matryca	Rozmiar: 1/1,8 cala kwadratowego; high-speed CMOS całkowita liczba pikseli: 6,6 megapikseli
Obiektyw/Ogniskowa	Od F2,7 (W) do 4,6 (T); f = 7,3 do 87,6mm (równoważnik od 36 do 432mm w formacie 35mm)
Zoom	12X zoom optyczny, 4X zoom cyfrowy
Regulacja ostrości	Autofocus, Makrofocus, Nieskończoność, Ręczne ustawianie ostrości
Przybliżony zakres ostrości (od powierzchni obiektywu)	Autofocus: od 40cm do $\infty$ (szeroki kąt) Makrofocus: od 5 do 50cm (szeroki kąt) Ręczne ustawianie ostrości: od 5cm do $\infty$
Ilość zdjęć wykonywanych na sekundę	Od 1 do 60 zdjęć na sekundę
Czułość	ISO 100, ISO 200, ISO 400, ISO 800, ISO 1600
Spust migawki	Ręczny spust na korpusie aparatu Zdalny spust migawki Samowyzwalacz 2s, 10s, potrójny

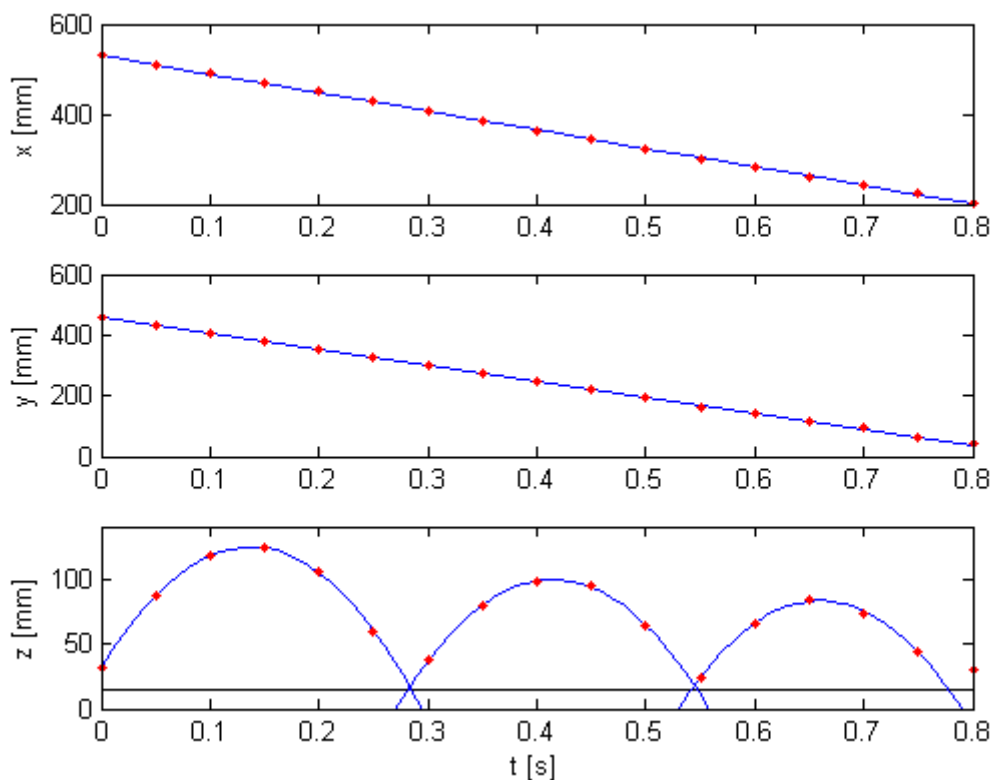
### **6.1 Analiza trajektorii odbijającej się piłki**

Pierwszy wykonany test sprawdza poprawność działania algorytmu i obliczeń przede wszystkim pod względem wizualnym. Do testu wykorzystana została piłeczka do ping-ponga, która odbijała się kilkakrotnie od płaszczyzny podstawy. Zarejestrowany pomiędzy odbiciami ruch jest zatem ruchem parabolicznym, powstającym ze złożenia jednostajnego ruchu prostoliniowego w rzucie na płaszczyznę  $XY$ , z ruchem pionowym przy niezrównoważonej sile grawitacji. Uzyskane dane z przetwarzania obrazów analizowane były osobno dla poszczególnych współrzędnych położenia w przestrzeni. Na wykresy naniesiono również wielomianowe przybliżenie uzyskanych punktów – dla osi  $X$  i  $Y$  jest to wielomian 1-stopnia, a dla osi  $Z$  są to trzy wielomiany 3-stopnia, które przybliżają trajektorię pomiędzy kolejnymi odbiciami od płaszczyzny podstawy.

Za miarę poprawności obliczeń można przyjąć błąd wyznaczenia pozycji, przy założeniu, że poprawna wartość wskazywana jest przez wielomian przybliżający. Innym wskaźnikiem jakości obliczeń jest punkt przecięcia się kolejnych wielomianów przybliżających współrzędną  $Z$  pomiędzy odbiciami – powinien znajdować się on na wysokości promienia piłki.

Przeprowadzone zostały dwa testy wg powyższego schematu, jeden przy świetle słonecznym, a drugi przy punktowym – ich wykresy przedstawione zostały odpowiednio na Rys. 25 i Rys. 26. Aby porównanie wyników pomiędzy testami było możliwe, fotografowana scena w obu przypadkach była podobna – odległość od kamery wynosiła około 1m, zdjęcia wykonywane były w rozdzielczości 2048x1536, a następnie skalowane w procesie przetwarzania w ten sam sposób. Pod względem wizualnym należy pozytywnie ocenić używany algorytm obliczeń – widoczne są wyraźne trendy trajektorii ruchu zgodne z oczekiwaniami. Należy jednak zwrócić uwagę, że wyniki zebrane podczas testu używającego światła żarówki, są wyraźnie mniej dokładne – występują większe odchylenia danych od interpolowanych funkcji. Analiza poszczególnych kroków procesu przetwarzania obrazu wskazuje na to, że główną przyczyną tej różnicy jest niska

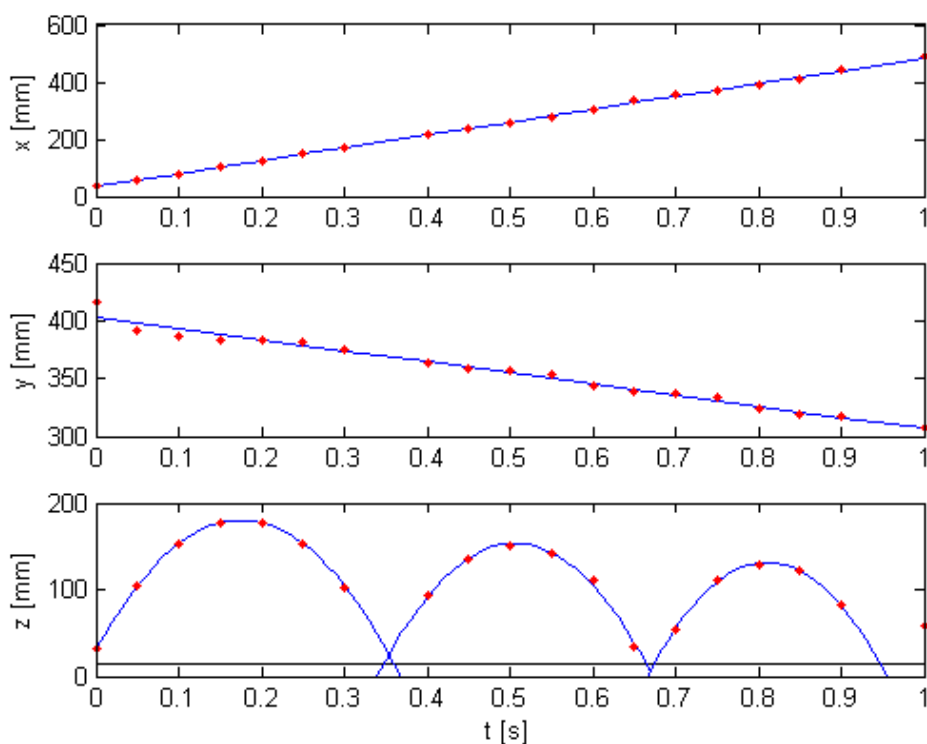
jakość zdjęcia i rozmycie obrazu spowodowane słabym oświetleniem sceny. Podczas testów używana była żarówka halogenowa o mocy 116W i strumieniu świetlnym 2135 lumenów.



Rys. 25: Trajektoria ruchu odbijającej się piłeczki (światło słoneczne)

Oświetlenie fotografowanej sceny było zatem znacznie słabsze niż przy naturalnym oświetleniu wykorzystanym przy pierwszym teście. Utrudniało to precyzyjne obliczenie środka szukanego na obrazie obiektu, co następnie wiązało się z błędami wyznaczenia pozycji.





Rys. 26: Trajektoria ruchu odbijającej się piłeczki (światło żarówki)

Możliwa jest również analiza otrzymanych wyników w oparciu o metody statystyczne. W Tabeli 2 przedstawiono odchylenie standardowe błędu wyznaczenia pozycji oraz średnia, minimalna i maksymalna wartość błędu dla poszczególnych współrzędnych  $X$ ,  $Y$  i  $Z$ . Dodatkowo zamieszczono wartość wysokości, na której, według interpolowanych wielomianów, nastąpiło odbicie. Wyniki tej analizy potwierdzają to, co zostało zaobserwowane wizualnie na wykresach – wyniki testu przeprowadzone przy świetle słonecznym są wyraźnie dokładniejsze. Mimo, że różnica błędu jest blisko trzykrotna, należy zauważyć, że błędy w obu przypadkach są niewielkie – największe wyznaczone odchylenie standardowe jest rzędu 10% rozmiaru obiektu.

---

Wykorzystanie cienia do odtworzenia pozycji obiektu w przestrzeni

---

*Tabela 2: Analiza statystyczna błędu wyznaczenia trajektorii ruchu*

Błąd wyznaczonego położenia [mm]	Światło punktowe			Światło słoneczne		
	<i>X</i>	<i>Y</i>	<i>Z</i>	<i>X</i>	<i>Y</i>	<i>Z</i>
Wartość minimalna	0,1041	0,3940	0,3571	0,3529	0,0172	0
Wartość maksymalna	6,8548	13,222	7,3357	3,5490	4,6275	2,1143
Wartość średnia	1,8872	2,8953	1,6132	1,4464	1,9043	0,5821
Mediana	1,3753	2,0318	0,8439	1,2451	1,4534	0,4571
Odchylenie standardowe	1,9255	3,1649	1,7437	0,7735	1,4802	0,5241
Wysokość odbicia	-8,5; 8,5			2; 2,5		

Wykonane testy pokazały, że dużym problemem jest łączenie się obrazu cienia i obiektu podczas uderzenia o podstawę. Nie udało się znaleźć rozwiązania tego problemu. Analizowano założenie, że w takiej sytuacji obiekt dotyka podstawy, czyli że szukany obiekt znajduje na znanej, niewielkiej wysokości (równej wysokości położenia środka obiektu leżącego na podstawie), która bezpośrednio przekłada się na współrzędną *Z* położenia. Następnie, znaleziony punkt wynikający z połączenia obrazu cienia i obiektu miał być rzutowany na płaszczyznę podstawy. W ten sposób miałyby być wyznaczone dwie pozostałe współrzędne. Należy jednak zauważyć niedokładności jakie wprowadza powyższy algorytm. Współrzędne środka znalezione na obrazie obiektu, po zrzutowaniu go na płaszczyznę podstawy, będą znacząco odbiegały od rzeczywistego położenia szukanego obiektu. Jest to spowodowane faktem, że środek analizowanej sumy obrazu obiektu i cienia jest wyraźnie przesunięty w górę obrazu. Wprowadza to duży błąd do wyznaczonej pozycji. Drugim elementem wprowadzającym niedokładności jest założenie, że połączenie obrazu obiektu i cienia oznacza, że obiekt dotyka płaszczyzny podstawy. W rzeczywistości może on znajdować się zdecydowanie wyżej. Zasięg tego błędu zależy będzie od ustawienia kamery, położenia światła i wielkości obiektu.

## 6.2 Wpływ kalibracji i błędu analizy obrazu na obliczaną pozycję obiektu

Jednym z najważniejszych parametrów, z punktu widzenia procesu technologicznego, wydaje się być powtarzalność wyników. Czynnikiem, który prawdopodobnie najbardziej wpływa na wyniki jest proces kalibracji systemu. W stworzonej aplikacji zdecydowano się na kalibrację ręczną. Użytkownik musi ręcznie wprowadzić odpowiednie punkty poprzez wskazanie ich na obrazie oraz przypisać im rzeczywiste współrzędne w przyjętym układzie odniesienia. Taka kalibracja może wprowadzać dwa źródła błędów – wpisanie niewłaściwych danych oraz niedokładnie wskazanie punktów na obrazie. Pierwsze z nich jest stosunkowo łatwe do wyeliminowania i nie wymaga dalszej analizy. Natomiast wskazanie odpowiednich punktów kalibracyjnych na fotografii może być obarczone dużym błędem ze względu na ich rozmazanie na obrazie. Cień, według którego kalibrowane jest światło, oraz wierzchołki znaczników kalibracyjnych są zawsze lekko zaokrąglone. Dodatkowo, zdjęcia są rejestrowane z małą rozdzielczością, więc również dyskretyzacja może wpływać na dokładność wprowadzanych danych. Tabela 3 przedstawia wyniki statystycznej analizy obliczonego położenia obiektu w przestrzeni dla jednego obrazu, lecz dla zmieniających się wskazań punktów kalibracyjnych. Zmiany te zawierały się zawsze w zakresie odległości 1 piksela (przyjmując sąsiedztwo Moora) od punktów przyjętych jako poprawne. W celu pokazania zależności wyników od rozdzielczości, obraz dwukrotnie przeskalowano, wyznaczono nowe punkty kalibracyjne przyjęte jako poprawne i ponownie przeprowadzono obliczenia.

Tabela 3: Analiza statystyczna błędu spowodowanego niedokładną kalibracją systemu

Rozdzielczość obrazu	1920x1440			1024x768			800x600		
	X	Y	Z	X	Y	Z	X	Y	Z
Wartość średnia błędu	0,89	2,07	1,03	1,56	3,80	1,96	2,06	4,83	2,46
Maksymalny błąd	4	6	5	6	12	9	8	14	11
Odchylenie standardowe	0,76	1,44	0,87	1,18	2,68	1,52	1,57	3,32	1,87

---

## Wykorzystanie cienia do odtworzenia pozycji obiektu w przestrzeni

---

Widoczny jest wpływ niedokładnej kalibracji na otrzymywane wyniki, jednak nie jest on duży. Dla obiektu wielkości ok 3,5cm, i rozdzielczości obrazu 1024x768 błąd średni wynosi 1,5-3,8mm, a maksymalny do 12mm. Systemy wizyjne używane w procesach przemysłowych są zazwyczaj kalibrowane jeden raz – podczas wdrażania linii lub wprowadzania nowego detalu, a następnie działają przez długi czas na uzyskanych ustawieniach. Jeżeli zatem niedokładny proces kalibracji nie wprowadza rażących błędów, a jedynie przesuwają układ współrzędnych o niewielki wektor, system wydaje się być stabilny i dobrze zdefiniowany.

Ważniejszą cechą jest powtarzalność działania skalibrowanego systemu. Elementem wprowadzającym błędy staje się niedokładny proces przetwarzania obrazu. Odblaski na obiekcie, przysłonięcia, czy źle dobrany próg binaryzacji mogą spowodować przesunięcie szukanego środka ciężkości znalezionego obiektu, co z kolei jest powodem błędnego wyznaczenia pozycji w przestrzeni. Tabela 4 przedstawia wyniki statystycznej analizy wyników uzyskanych dla sztucznie wygenerowanych błędów wyznaczenia pozycji obiektów na obrazie. Wprowadzany błąd przekraczał 25% wielkości obiektu, którego średnica wynosiła 30,48mm, a używana rozdzielczość to 1024x768.

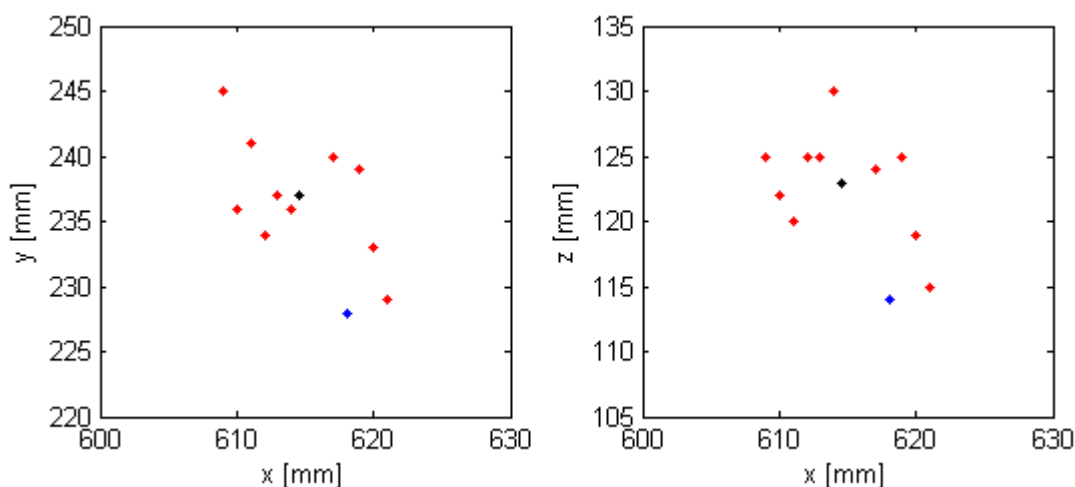
*Tabela 4: Analiza statystyczna błędów związanych z niedokładnym przetwarzaniem obrazu*

	X [mm]	Y [mm]	Z [mm]
Wartość średnia	469,107	379,799	123,585
Mediana	469	380	124
Wartość maksymalna	476	391	131
Wartość minimalna	463	368	116
Maksymalna odchyłka dodatnia	7	11	7
Maksymalna odchyłka ujemna	6	12	8
Odchylenie standardowe	2,5919	4.9834	3,2426

Błąd wynikający z niedokładnego wyznaczenia środka obiektu jest zatem porównywalny z opisanym wcześniej błędem niedokładnej kalibracji. Jest on jednak dużo większym problemem, ponieważ może znacząco wpływać na działanie procesu produkcyjnego. Kalibracja jest wykonywana rzadko i po każdej jej zmianie możliwa jest ponowna regulacja pozostałych części systemu. Ten błąd natomiast jest niezależny od kalibracji i pojawiać się może przy każdym kolejnym analizowanym obrazie. Z tego względu wydaje się być konieczna bardziej zaawansowana obróbka obrazu i jego segmentacja w taki sposób, aby znalezienie środka obiektu było jak najdokładniejsze.

### **6.3 Wpływ zmiany położenia aparatu**

Innym testem pozwalającym sprawdzić powtarzalność otrzymywanego rozwiązania, jest sfotografowanie tej samej nieruchomej sceny z wielu różnych perspektyw. Należy zauważyć, że na ten test mają wpływ również błędy opisane w poprzednim podrozdziale – błędy kalibracji, ponieważ dla każdego kolejnego położenia aparatu system musiał zostać ponownie skalibrowany. Test ten wykonany został jedynie dla światła punktowego (żarówki) dzięki czemu stała była nie tylko pozycja obiektu, ale również punkt światła oraz położenie cienia na płaszczyźnie podstawy. Na Rys. 27 przedstawione zostały wyniki obliczeń położenia obiektu przy dziesięciu różnych położeniach aparatu. Kolorem niebieskim zaznaczony został punkt ręcznie zmierzonego położenia, a kolorem czarnym wartość średnia wszystkich obliczonych pozycji.



Rys. 27: Wpływ zmiany położenia aparatu na wyznaczoną pozycję obiektu

Różnice pomiędzy wyznaczonymi pozycjami są podobne do uzyskanych w teście wpływu kalibracji. Oznaczać to może, że wszystkie te błędy mogą być spowodowane niedokładną kalibracją. Maksymalne odchylenie wynosi 8mm, a odchylenie standardowe nie przekracza 5mm. Należy jednak zwrócić uwagę na różnicę pomiędzy zmierzonym a wyznaczonym położeniem. W idealnym przypadku, wartość zmierzona powinna być zbliżona do średniej wartości obliczeń. Różnica ta jest jednak niewielka. Biorąc pod uwagę kalibrację systemu i możliwość dodania stałej wartości wyrównującej wyniki obliczeń, w celu zbliżenia się do rzeczywistości, błąd ten można przyjąć jako nieistotny. Błąd ten może wynikać również z niedokładnie wykonanego pomiaru przy użyciu ręcznej miarki.

#### **6.4 Czas przetwarzania obrazu**

Bardzo ważnym problemem w systemach wizyjnych jest czas potrzebny na analizę obrazu. Operacje przetwarzania obrazu są zazwyczaj bardzo wymagające obliczeniowo.

Zaprezentowany algorytm pracuje stosunkowo szybko. Okazuje się, że najbardziej czasochłonnym krokiem jest wczytanie obrazu z dysku twardego. Zdjęcia wykonane były z rozdzielczością 2048x1536 pikseli. Na komputerze, na którym przeprowadzane były testy, czas wczytania takiego obrazu z pliku

wynosił pomiędzy 60 a 70 milisekund. Jest to stosunkowo długo, biorąc pod uwagę fakt, że cały czas przetworzenia takiego zdjęcia, aż do uzyskania wyniku wynosił pomiędzy 31 a 36 milisekund. Jeżeli natomiast obraz był dodatkowo skalowany, do rozdzielczości 1024x768, czas przetwarzania spadał do 7-8 milisekund (wliczając czas skalowania).

Biorąc pod uwagę rozdzielczości używane zazwyczaj w przemysłowych systemach wizyjnych, czas przetwarzania obrazu używając zaproponowanego algorytmu jest niski. Problemem może nawet okazać się czas akwizycji obrazu, ze względu na stosunkowo długie naświetlenie potrzebne przy niezbyt mocnym i nierozproszonym oświetleniu sceny. Innym ograniczeniem czasowym może być nawet czas potrzebny na przesłanie zarejestrowanego obrazu do jednostki obliczeniowej.

## 7 Kierunki dalszego rozwoju pracy

Założeniem stworzonej aplikacji było przede wszystkim wypróbowanie działania algorytmu i umożliwienie oceny zaproponowanej metody wyznaczenia pozycji obiektu w przestrzeni na podstawie jednego obrazu. Aby mogła ona być wykorzystana w procesie produkcyjnym, musiałaby zostać zaprojektowana jako system czasu rzeczywistego. Dodatkowo, należałoby dodać moduł pobierający obraz z pracującej kamery, a nie tylko z dysku twardego lub innego nośnika danych, co umożliwia stworzona aplikacja.

Rozważania nad poprawieniem dokładności systemu należałoby skierować przede wszystkim w stronę badań nad lepszą analizą obrazu. Błąd wynikający z nieprecyzyjnego wyznaczenia środków obiektów znajdujących się na obrazie przekładał się w znaczny sposób na otrzymywane wyniki. Zaproponowane w podrozdziale 4.3 rozwiązania są jedynie propozycją wymagającą zweryfikowania. Zarówno wyszukiwanie obiektu za pomocą jego barwy, jak i użycie algorytmów generacji tła, powinny przynieść poprawę dokładności, a przede wszystkim powtarzalności uzyskiwanych wyników. Należy również przeanalizować dodanie algorytmu pozwalającego ocenić poprawność znalezionych środków obiektu i cienia na obrazie. Aktualnie zaimplementowany algorytm, w przypadku niepoprawnego rozpoznania obrazu, może zaklasyfikować szum jako szukany obiekt. Próby wyeliminowania tego efektu należałoby rozpocząć od oceny rozmiaru znalezionego po erozji obiektu.

Kolejnym pomysłem, który wydaje się być wartym przetestowania, jest użycie kilku źródeł światła. Większa ilość cieni, czyli większa ilość obrazów szukanego obiektu może przyczynić się do poprawy dokładności obliczeń. Obiekt, w takim przypadku, może nawet nie być widoczny w kadrze – wyznaczenie pozycji mogłoby odbywać się jedynie na podstawie cieni. Dodatkowo, polepszyłyby się oświetlenie sceny.

W niniejszej pracy, analizowano jedynie scenę, gdzie cień zawsze padał



---

### Wykorzystanie cienia do odtworzenia pozycji obiektu w przestrzeni

---

na płaską poziomą powierzchnię. W ogólniejszym, lecz trudniejszym przypadku, należałoby umożliwić wprowadzenie kilku płaszczyzn, na których może znaleźć się cień lub nawet stworzenie odpowiedniej trójwymiarowej mapy powierzchni.

Elementem ułatwiającym pracę z aplikacją byłoby też zautomatyzowanie procesu kalibracji poprzez stworzenie wzorca kalibracyjnego i programowe wyznaczanie jego punktów charakterystycznych. Zadaniem operatora byłoby jedynie odpowiednie umiejscowienie wzorca w widocznym punkcie sceny i ewentualne wprowadzenie wartości korygujących odsunięcie i obrót wzorca względem przyjętego układu współrzędnych.

## Podsumowanie

W pracy zaproponowane zostało wyznaczanie pozycji obiektu w przestrzeni w oparciu o jeden obraz, na którym widoczny jest obiekt oraz jego cień. Występuje tutaj redukcja ilości zdjęć potrzebnych do analizy w stosunku do tradycyjnych systemów stereowizyjnych. Testy pokazały dobrą dokładność i powtarzalność obliczeniową metody.

Największe błędy spowodowane były niedokładnym wyznaczaniem środka obiektu i jego cienia na zarejestrowanym obrazie. Przyczyną okazało się być niedostateczne oświetlenie sceny. Aby możliwe było użycie zaproponowanej metody wyznaczania pozycji obiektu, używane światło musi spełniać pewne wymogi. Jednym z nich jest punktowa natura źródła światła. Pozwala ona na uzyskanie wyraźnego cienia na płaszczyźnie podstawy. Z drugiej strony, jedno mocne, punktowe światło jest przyczyną pojawienia się półcieni, utrudniających odnajdywanie obiektów. Przy wykorzystywaniu światła sztucznego problemem jest też dostateczne doświetlenie całej sceny. Konieczne jest zastosowanie źródła światła o dużej mocy i nie jest możliwe używanie kloszy odbijających światło w stronę sceny, ponieważ cień staje się rozmyty.

Światło słoneczne wydaje się zmniejszać część trudności związanych z obróbką. Zdjęcia w pogodny dzień są ostre, dobrze doświetlone, a cień jest wyraźny. Niestety światło słoneczne nie zawsze jest dostępne, więc jego zastosowanie w przemyśle jest ograniczone.

Stworzona aplikacja pozwoliła na pomyślne przetestowanie zaproponowanego systemu. Przed jej użyciem w procesie produkcyjnym musiałaby jednak zostać nieznacznie zmieniona i ulepszona, przede wszystkim w celu nadania jej cech systemu czasu rzeczywistego. Stanowi ona jednak dobrą podstawę do dalszych badań nad zaproponowaną metodą lokalizacji obiektu w przestrzeni. Mogłaby być również wykorzystywana w aktualnej wersji do prototypowania systemów np. poprzez znajdowanie optymalnych ustawień

---

## Wykorzystanie cienia do odtworzenia pozycji obiektu w przestrzeni

---

procesów przetwarzania obrazu, podczas gdy na linii produkcyjnej obliczenia mogłyby być wykonywane na mikrokontrolerach lub układach FPGA.

Metoda wykorzystująca jedną kamerę wydaje się być atrakcyjną alternatywą w stosunku do tradycyjnych systemów stereowizyjnych. Powodem jest zmniejszenie kosztów wdrożenia takiego systemu, ze względu na użycie tylko jednej kamery, przy zachowaniu dobrej dokładności wyznaczonej pozycji.

## Literatura

- [1]: Cyganek B., Komputerowe przetwarzanie obrazów trójwymiarowych, 2002.
- [2]: Santos P., Dee H. M., Fenelon V., Notes on a qualitative theory of shadows, 2008.
- [3]: Fuksa S., Byrski W., "Czteropunktowa metoda identyfikacji transformacji stereowizyjnej", Automatyka / Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie, 2005, T. 9, Z. 3.
- [4]: OpenCV 2.4.5.0 documentation, <http://docs.opencv.org/>, 2013.
- [5]: Tadeusiewicz R., Korohoda P., Komputerowa analiza i przetwarzanie obrazów, 1997.
- [6]: Kryjak T., Analiza i testowanie algorytmów generacji tła na potrzeby systemu monitoringu przestrzeni publicznej, Automatyka / Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie, 2011, T. 15, Z. 3.
- [7]: Jerzy Grębosz, Symfonia C ++ Standard , 2000.
- [8]: "Qt Documentation Archives", <http://doc.qt.digia.com/>, 2013.
- [9]: Blanchette J., Summerfield M., C++ GUI Programming with Qt 4, 2006.
- [10]: Robert Laganniere, OpenCV 2 Computer Vision Application Programming Cookbook, 2011.
- [11]: Armadillo, C++ linear algebra library, <http://arma.sourceforge.net/>, 2013.
- [12]: Documentation Center, [www.mathworks.com/help/matlab/](http://www.mathworks.com/help/matlab/), 2013.
- [13]: CASIO, Aparat Cyfrowy EX-F1, Podręcznik użytkownika, 2008.

## **Załączniki**

Do pracy na płycie CD załączone zostały:

1. Aplikacja wraz z bibliotekami potrzebnymi do jej uruchomienia.
2. Pliki z zapisanymi ustawieniami przetwarzania obrazu poszczególnych wykonywanych testów.
3. Skrypty programu MATLAB używane do analizy wyników i rysowania wykresów.
4. Pliki kodu źródłowego aplikacji (projekt stworzony w programie QT Creator).
5. Część wykonanych zdjęć używanych podczas testów.