

Przemysław KOWALSKI, Krzysztof SKABEK  
Instytut Informatyki Teoretycznej i Stosowanej PAN

## PRZETWARZANIE INFORMACJI WIZYJNEJ W KOMPUTEROWYM SYSTEMIE Z MOBILNĄ GŁOWICĄ STEREOWIZYJNĄ

**Streszczenie.** Do badań wykorzystano sterowany zdalnie z komputera niewielki pojazd gasienicowy z wysięgnikiem, na którym umieszczono głowicę stereowidzącą. Zadaniem systemu jest utworzenie reprezentacji przestrzennej obiektów obserwowanej sceny 3D na podstawie sekwencji obrazów stereoskopowych pochodzących z różnych punktów widokowych rozmieszczonych wokół sceny statycznej. Do reprezentacji obiektów sceny wykorzystano grafy struktury. Opisano także grafowe algorytmy rekonstrukcji sceny 3D.

## PROCESSING OF VISUAL INFORMATION IN THE COMPUTER SYSTEM WITH THE MOBILE STEREO HEAD

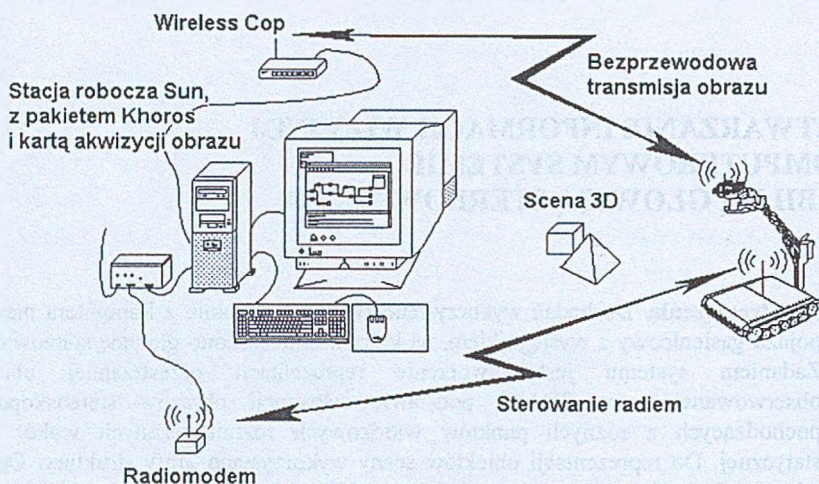
**Summary.** As a hardware platform we used a research vehicle with a stereovision head controlled by a remote computer. The task of the presented system is construction of a complete 3D scene representation on the basis of stereo sequences from different points of view placed around the static scene. Structure graphs were used to represent objects of 3D scene. Graph algorithms for 3D scene reconstruction were described.

### 1. Platforma badawcza

Podstawą akwizycji obrazów w systemie jest aktywna głowica stereowidząca osadzona na niewielkim pojeździe gasienicowym z wysięgnikiem. Sterowanie pojazdem odbywa się za pomocą pięciu silników elektrycznych pozwalających na jazdę do przodu, do tyłu, skrzyty,

podnoszenie i opuszczanie wysięgnika z głowicą stereowidzącą, a także obrót oraz podnoszenie i opuszczanie głowicy.

Przyjęto model, w którym dana jest scena statyczna (nieruchome obiekty, stacjonarne oświetlenie itp.), natomiast kamery mają możliwość obrotu wokół określonego punktu centralnego sceny.



Rys. 1. Schemat systemu aktywnego stereowidzenia

Fig. 1. Scheme of the active stereovision system

## 2. Przesyłanie informacji wizyjnej w systemie

Aktywna głowica stereowidząca jest połączona z komputerem (stacją roboczą Sun) za pośrednictwem dwóch łączy bezprzewodowych. Pierwsze z nich służy do transmisji obrazów, a drugie do transmisji poleceń sterujących [1, 2].

Mimo posiadania przez głowicę dwóch kamer, nie istnieje możliwość równoczesnej transmisji obrazów z obu kamer. Wynika to z faktu wykorzystania pojedynczego łącza do transmisji obrazów. Zastosowany zestaw „Wireless Cop”, składający się z odbiornika i nadajnika sygnałów audio-video, jest w zasadzie przeznaczony dla zastosowań telewizyjnej przemysłowej, w których może pracować z wykorzystaniem wielu różnych kanałów. Kanały w odbiorniku są jednak przełączane ręcznie, co uniemożliwia ich programowe przełączanie; podobnie wyboru kanału w nadajniku dokonuje się zwierając odpowiednie zworki. Odbiornik dysponuje pojedynczymi wyjściami dla sygnału video i audio.

Karta akwizycji obrazu – Sun Video Ultra Plus, dysponuje dwoma wejściami sygnału video. Zaleta ta nie może być jednak wykorzystana, z uwagi na użycie pojedynczego zestawu „Wireless Cop”, z jednym kanałem wybranym przed rozpoczęciem pracy systemu.

Karta akwizycji może pracować z obrazami w formatach NTSC (640x480) i PAL (768x576). W ramach projektu zastosowano dwie czarno-białe kamery CCD i system PAL. Polecenie akwizycji obrazów wbudowano w procedurę sterowania aktywną głowicą stereowidzącą – może zostać wywołane jako jedno z poleceń języka skryptowego. W procesie akwizycji obrazów wykorzystuje się funkcje należące do biblioteki XIL.

Kamery są przełączane zdalnie przy wykorzystaniu łącza bezprzewodowego dla poleceń sterujących. Rozwiązanie takie uniemożliwia pracę w czasie rzeczywistym, ze względu na opóźnienia w przełączaniu kamer, mające wpływ na czas akwizycji.

### 3. Sterowanie platformą wizyjną

Kanał sterowania zaprojektowano opierając się na standardzie RS 232c. Aby umożliwić autonomiczną pracę głowicy do transmisji użyto dwa radiodemody, pracujące na częstotliwości 433,92 MHz. Wykorzystano tryb transmisji 8-bitowej, z bitem parzystości i jednym bitem stopu; o prędkości 1200 bitów na sekundę.

Kanał sterowania pracuje w trybie half-duplex – wynika to z wykorzystania protokołu BNET (opracowanego w firmie SCS), dla transmisji poleceń.

Lista poleceń sterujących aktywną głowicą stereowizyjną obejmuje:

- ruch pojazdu (Forward, Backward, Left, Right, ArmUp, ArmDown, HeadUp, HeadDown, HeadLeft, HeadRight);
- włączanie i wyłączanie kamer, akwizycję obrazu (GetImage, GetStereo, GetLeft, GetRight).

Pełniejszy opis języka skryptowego, wykorzystywanego jako interfejs do wydawania poleceń aktywnej głowicy stereowidzącej; a odpowiadający dostępnym poleceniom w celu sterowania głowicą, zawarto w [1, 2].

Polecenia przesyłane są do sterownika aktywnej głowicy stereowidzącej w postaci scii wpisów do rejestrów. Przykładowo akwizycja pary stereoobrazów przebiega w kilku etapach:

1. Włączenie kamery numer 2 (lewej).
2. Włączenie kamery, wymaga zapisania w rejestrze sterownika kilku kolejnych wartości:
  - wstawienie do rejestru sterownika numer dwa numeru kamery (przesyłana ramka SetExt[3]);

- wstawienia do rejestru sterownika numer jeden wartości '5' (odpowiada poleceniu „włącz kamerę”);
  - wstawienie do rejestru sterownika numer zero, wartości '5' (odpowiada zgłoszeniu poprawności połączenia);
  - uruchomienie programu sterownika (przesyłana ramka BasRun);
  - wysłanie do rejestru sterownika numer trzy wartości '-1' (odpowiada ustawieniu flagi gotowości, czyli wywołaniu uruchomienia procedury).
3. Akwizycja obrazu;
  4. Włączenie kamery numer 1 (prawej).
  5. Akwizycja obrazu.
  6. Wyłączenie kamery numer 1 (prawej). Wyłączenie przebiega podobnie do włączenia, z wyjątkiem ustawienia rejestru numer jeden – dla polecenia wyłączenia kamery jest on ustawiany na wartość '6'.

Jak łatwo wyliczyć, wykonanie akwizycji pary stereobrazów wymaga przesłania do sterownika 15 ramek protokołu BNET. Odbiór każdej ramki jest potwierdzany przez sterownik (odbierając ramkę SetExt sterownik odpowiada sygnalizując ewentualny błąd – ramka BasErr; BasRun odpowiada BasStat).

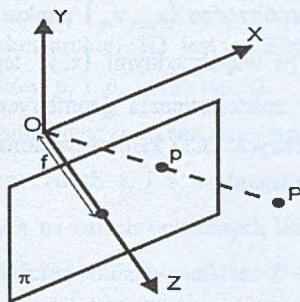
Typowa ramka w protokole BNET składa się z adresu (1 bajt), numeru ramki (2 bity) i kodu polecenia (6 bitów); przesyłanych danych oraz 2 bajtów sumy kontrolnej CRC. Ramka typu SetExt (ustawianie zawartości rejestru) zawiera 4 bajty danych, liczy więc w sumie 8 bajtów; ramka BasRun zawiera 33 bajty danych; ramka BasErr zawiera jeden bajt danych; a ramka BasStat 3 bajty danych. Łącznie jednemu włączeniu/wyłączeniu kamer odpowiada przesłanie 96 bajtów, czyli uwzględniając wyłącznie przesyłanie danych, realizacja takiego polecenia nie może trwać krócej niż 0,8 s; a cały proces (3 przełączenie) – 2,4 s. W wyliczeniu nie uwzględniono ewentualnych retransmisji, wynikających z możliwych zakłóceń w komunikacji.

Oprócz opóźnień wynikających z czasu transmisji poleceń przełączenia kamer, pojawiają się również opóźnienia wynikające z samego czasu realizacji akwizycji obrazu przez kartę Sun Video Ultra Plus, oraz późniejszego przetwarzania uzyskanych obrazów. Wynika stąd, że nie można dokonywać analizy obrazów uzyskanych przez aktywną głowicę stereowidzącą na bieżąco w czasie akwizycji – wymusza to przyjęcie strategii nawigacji mającej postać serii poruszeń głowicą występujących naprzemiennie z akwizycjami obrazów i ich analizą.

## 4. Przetwarzanie rastrowych obrazów stereo

### 4.1. Model rzutowania

Obraz rejestrowany przez pojedynczą kamerę jest rzutem perspektywicznym obserwowanej sceny przestrzennej. Model przekształcenia perspektywicznego zaprezentowano na rys. 2. Model składa się z płaszczyzny obrazu  $\pi$  oraz punktu przestrzennego  $O$  będącego środkiem rzutowania. Odległość pomiędzy  $\pi$  i  $O$  jest długością ogniskowej  $f$ . Prosta przechodząca przez punkt  $O$  oraz prostopadła do płaszczyzny  $\pi$  nazywa się osią optyczną. Oś optyczna przecina płaszczyznę  $\pi$  w punkcie  $O$  zwanym punktem centralnym.



Rys. 2. Model perspektywiczny kamery

Fig. 2. Model of perspective camera

Punkt  $p = [x, y, z]^T$  jest obrazem punktu  $P = [X, Y, Z]^T$  na płaszczyźnie  $\pi$ . W układzie odniesienia kamery dla modelu perspektywicznego prawdziwe jest następujące przekształcenie:

$$x = f \frac{X}{Z}, \quad y = f \frac{Y}{Z}, \quad z = f. \quad (1)$$

Algorytmy rekonstrukcji sceny przestrzennej oraz obliczania położenia obiektów w przestrzeni wymagają równań wiążących współrzędne punktów w przestrzeni 3D ze współrzędnymi ich odpowiedników na płaszczyźnie obrazu. W celu powiązania różnych układów odniesienia (układ globalny, układ kamery, układ obrazu) konieczne jest oszacowanie parametrów wewnętrznych i zewnętrznych kamery.

Parametry zewnętrzne definiują położenie i orientację układu odniesienia kamery w stosunku do globalnego układu odniesienia. Podstawowe parametry tego przekształcenia to:  $T$  – wektor translacji opisujący względne położenie środków dwóch układów odniesienia,  $R$  – macierz obrotu o wymiarze  $3 \times 3$ , macierz ortogonalna powodująca nałożenie

odpowiadających sobie osi układów. Ortogonalność relacji zmniejsza liczbę stopni swobody do trzech, co odpowiada kątom obrotu:  $\alpha, \beta, \gamma$ .

Relacja pomiędzy współrzędnymi punktu  $P$  w globalnym układzie współrzędnych  $P_w$  oraz układzie współrzędnych kamery  $P_C$  jest następująca:

$$P_C = R(P_w - T). \quad (2)$$

Parametry wewnętrzne charakteryzują optyczne, geometryczne i cyfrowe cechy kamery. Do modelu perspektywicznego kamery wymagane są trzy rodzaje parametrów wewnętrznych określające: rzut perspektywiczny, dla którego jedynym parametrem jest długość ogniskowej  $f$ , przekształcenie pomiędzy układem współrzędnych kamery a współrzędnymi pikseli, zniekształcenia geometryczne toru optycznego.

Należy powiązać ze sobą współrzędne  $(x_{im}, y_{im})$  punktu obrazu w układzie odniesienia obrazu wyrażone w pikselach ze współrzędnymi  $(x, y)$  tego samego punktu w układzie odniesienia kamery. Pomijając zniekształcenia geometryczne wprowadzane przez układ optyczny oraz zakładając, że matryca CCD kamery wykonana jest jako prostokątna siatka elementów fotoczułych, mamy równania:

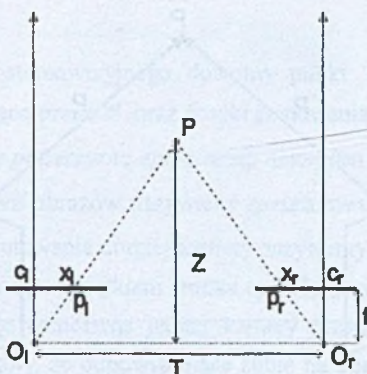
$$\begin{cases} x = -(x_{im} - o_x) s_x \\ y = -(y_{im} - o_y) s_y \end{cases}, \quad (3)$$

gdzie:  $(o_x, o_y)$  – współrzędne środka obrazu w pikselach,  $(s_x, s_y)$  – efektywny rozmiar piksela w  $mm$ , odpowiednio w kierunku poziomym i pionowym.

#### 4.2. Zagadnienia stereowidzenia

Stereowizja zajmuje się pozyskiwaniem informacji o strukturze 3D oraz odległości w scenie na podstawie dwóch lub więcej obrazów pochodzących z różnych punktów widzenia [7]. Za pomocą systemu stereowizyjnego można rozwiązać zarówno zadanie ustalenia odpowiedniości, jak również rekonstrukcji 3D. Ustalenie odpowiedniości polega na określeniu, które punkty z jednego obrazu odpowiadają punktom na drugim oraz odrzuceniu punktów przesłoniętych. Mając daną pewną liczbę odpowiadających sobie punktów w dwóch obrazach oraz informacje o położeniu obrazów w układzie globalnym, możemy odtworzyć położenie i strukturę obserwowanych obiektów. Ustalone względne odległości pomiędzy odpowiadającymi sobie elementami obrazów tworzą *mapę rozbieżności*<sup>1</sup>. Jeśli znana jest geometria systemu stereowizyjnego, to mapę rozbieżności można przekonwertować do postaci mapy przestrzennej obserwowanej sceny.

<sup>1</sup> ang. disparity map



Rys. 3. Obliczanie głębi w układzie stereowizyjnym  
Fig. 3. Depth estimation in a stereo system

Zadaniem podstawowym rekonstrukcji 3D jest odnalezienie położenia w przestrzeni punktu  $P$  na podstawie jego rzutów  $p_l$  i  $p_r$  (patrz rys. 3).

Odległość  $T$  wyznaczono pomiędzy środkami rzutowania  $O_l$  i  $O_r$  wzdłuż tzw. linii bazowej<sup>1</sup> systemu stereowizyjnego. Niech  $x_l$  i  $x_r$  będą współrzędnymi  $p_l$  i  $p_r$  w odniesieniu do punktów  $c_l$  i  $c_r$  leżących na osiach optycznych kamer,  $f$  jest wspólną długością ogniskowej kamer, zaś  $Z$  odległością pomiędzy punktem  $P$  a linią bazową.

Dla opisanego powyżej układu prawdziwa jest zależność:

$$Z = f \frac{T}{d}, \quad (4)$$

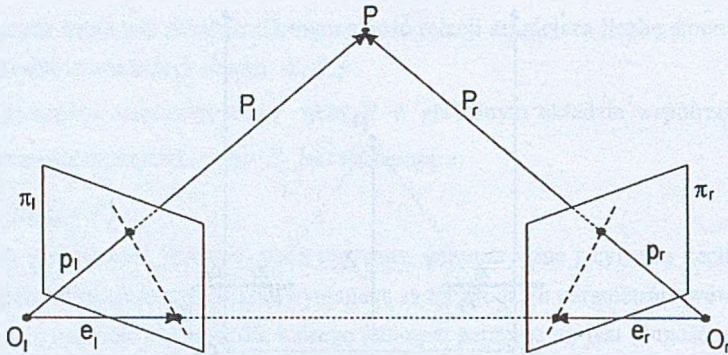
gdzie:  $d = x_r - x_l$  określa rozbieżność<sup>2</sup>, czyli różnicę względnych odległości od osi optycznych pomiędzy odpowiadającymi sobie punktami dwóch obrazów.

Z powyższego równania widać, że głębokość w scenie jest odwrotnie proporcjonalna do zmierzonej rozbieżności oraz wprost proporcjonalna do długości ogniskowej i odległości pomiędzy kamerami.

Powyżej określone parametry możemy podzielić na dwa rodzaje: wewnętrzne i zewnętrzne. Parametry wewnętrzne charakteryzują odwzorowanie przekształcające punkt obrazu ze współrzędnych w układzie kamery do współrzędnych obrazowych dla każdej z kamer. Parametry te są identyczne z tymi wprowadzonymi wcześniej dla pojedynczej kamery. Parametry zewnętrzne opisują względną pozycję oraz orientację dwóch kamer, czyli przekształcenie sztywne (obrót i translacja) powodujące nałożenie na siebie układów odniesienia dwóch kamer.

<sup>1</sup> ang. baseline

<sup>2</sup> ang. disparity



Rys. 4. Geometria epipolarna  
Fig. 4. The epipolar geometry

W przypadku gdy parametry wewnętrzne i zewnętrzne są nieznanne, zagadnienie rekonstrukcji rozpoczyna się zwykle od dokonania kalibracji<sup>1</sup>. Jednakże może się okazać, że system stereowizyjny pozyskuje dużą ilość informacji o obserwowanej scenie bez wcześniejszej kalibracji. Mamy wtedy do czynienia z układem stereowizyjnym nieskalibrowanym.

### 4.3. Geometria epipolarna

Ważnym założeniem dla układu stereowizyjnego jest stosowanie geometrii epipolarniej. Na rys. 4 przedstawiono dwie kamery perspektywiczne, ich środki rzutowania:  $O_l$ ,  $O_r$  i płaszczyzny obrazów:  $\pi_l$ ,  $\pi_r$ . Długości ogniskowych wynoszą:  $f_l$  i  $f_r$ . Każda kamera umieszczona jest we własnym układzie odniesienia, którego środek jest w środku rzutowania, a oś  $Z$  jest osią optyczną. Wektory  $P_l = [X_l, Y_l, Z_l]$  oraz  $P_r = [X_r, Y_r, Z_r]$  odnoszą się do tego samego punktu w przestrzeni  $P$ , który można potraktować jako wektor odpowiednio w prawym i lewym układzie odniesienia kamer. Wektory  $p_l = [x_l, y_l, z_l]$  oraz  $p_r = [x_r, y_r, z_r]$  odpowiadają rzutom  $P$  na, odpowiednio, lewą i prawą płaszczyznę obrazu i wyrażone są w odpowiednim układzie odniesienia. Dla wszystkich punktów obrazu prawdą jest odpowiednio:  $z_l = f_l$  lub  $z_r = f_r$ .

Układy odniesienia lewej i prawej kamery są powiązane za pomocą parametrów zewnętrznych układu stereowizyjnego. Definiują one sztywne przekształcenie przestrzeni 3D za pomocą wektora przesunięcia  $T = (O_r - O_l)$  oraz macierzy rotacji  $R$ . Dla punktu  $P$  można określić relację pomiędzy  $P_l$  i  $P_r$ :

<sup>1</sup> Kalibracja systemu stereowizyjnego to proces wyznaczania jego parametrów wewnętrznych i zewnętrznych.



$$P_r = R(P_l - T). \quad (5)$$

Dla zadanego układu stereowizyjnego dowolny punkt  $P$  z przestrzeni 3D określa płaszczyznę  $\pi_p$ , przechodzącą przez  $P$  oraz środki rzutowania obydwu kamer (patrz rys. 4). Płaszczyznę  $\pi_p$  nazywamy *płaszczyzną epipolarną*, natomiast proste przecięcia płaszczyzny  $\pi_p$  z płaszczyznami obydwu obrazów nazywamy *sprzężonymi liniami epipolarnymi*. Obraz w jednej kamerze środka rzutowania drugiej kamery nazywamy *środkiem epipolarnym*<sup>1</sup>.

Przez każdy punkt obrazu, z wyjątkiem środka epipolarnego, przechodzi tylko jedna linia epipolarna. Wszystkie linie epipolarne jednej kamery przechodzą jej środek epipolarny. Założenie epipolarności mówi, że odpowiadające sobie na obrazach punkty muszą leżeć na sprzężonych liniach epipolarnych.

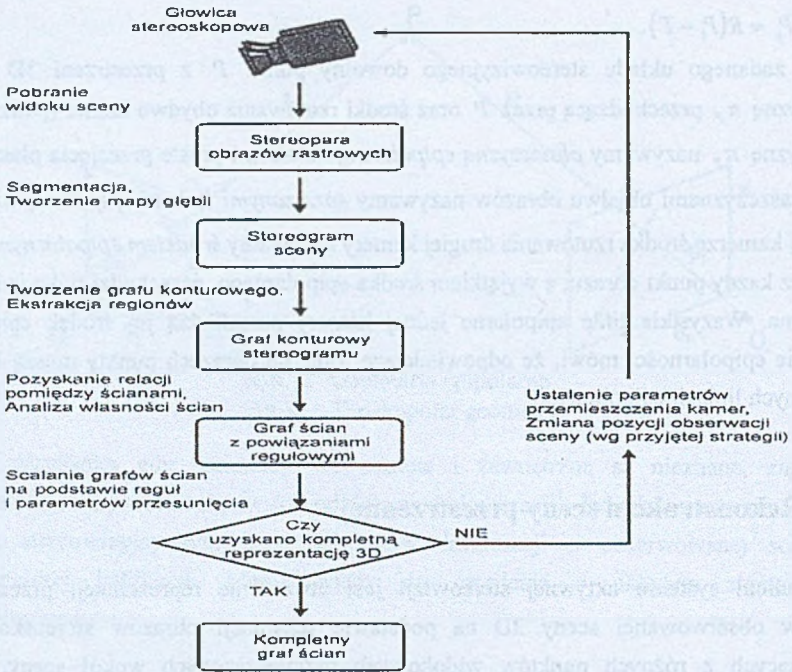
## 5. Rekonstrukcja sceny przestrzennej

Zadaniem systemu aktywnej stereowizji jest utworzenie reprezentacji przestrzennej obiektów obserwowanej sceny 3D na podstawie sekwencji obrazów stereoskopowych pochodzących z różnych punktów widokowych rozmieszczonych wokół sceny. Proces rekonstrukcji sceny zaprezentowany na rys. 5 składa się z kilku etapów: 1) akwizycja rastrowych stereoskopowych widoków sceny, 2) segmentacja i tworzenie mapy głębi, 3) tworzenie grafu konturu, 4) ekstrakcja regionów, tworzenie grafu ścian, 5) analiza właściwości ścian oraz relacji pomiędzy ścianami, 6) scalanie grafu aktualnego widoku z modelowym grafem ścian, 7) sprawdzenie kompletności reprezentacji.

Na poszczególnych etapach procesu rekonstrukcji sceny dane zmieniają swoją reprezentację, poczynając od stereopary obrazów rastrowych tworzącej po uwzględnieniu mapy głębi stereogram sceny (jego wymiar możemy określić jako  $2\frac{1}{2}D$ ). W wyniku segmentacji można wyodrębnić ze stereogramu graf konturu, a po ekstrakcji regionów tworzony jest graf ścian wzbogacony dodatkowo właściwościami ścian oraz relacjami pomiędzy ścianami. Ostatecznie tworzony jest kompletny graf ścian, a wraz z nim odpowiedni graf konturu.

---

<sup>1</sup> ang. epipole



Rys. 5. Etapy rekonstrukcji sceny 3D

Fig. 5. Stages of the 3D scene reconstruction

Na etapie przetwarzania wstępnego para obrazów rastrowych podlega segmentacji mającej na celu wyodrębnienie konturów oraz wierzchołków [4]. Uzyskane w ten sposób kontury podlegają następnie pocienianiu oraz likwidacji nieciągłości. Dodatkowo tworzona jest mapa głębi dla punktów krytycznych, czyli wierzchołków i punktów przecięcia każdej stereopary. Wyznaczanie odpowiedności oraz konstrukcja mapy głębi wykonywane są metodą konturów aktywnych opisaną w artykule [5].

W rezultacie konstruowane są grafy struktury dla pozyskanego widoku sceny.

## 6. Grafowa reprezentacja struktur przestrzennych

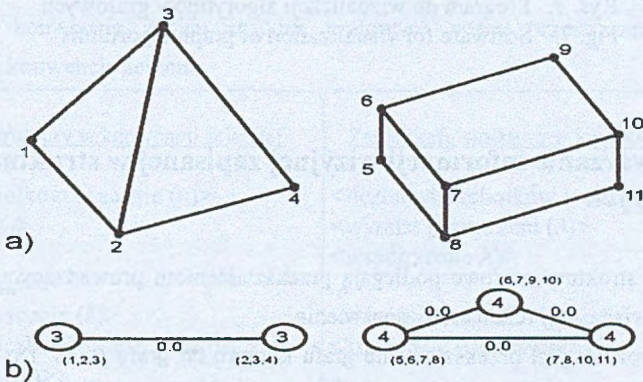
Pojęcie struktury obrazu można zdefiniować jako charakterystyczny dla każdego elementu obrazu sposób rozmieszczenia tych elementów wraz z zespołem relacji pomiędzy nimi.

Naturalną reprezentacją tak określonej struktury jest odpowiednio skonstruowany graf zwany grafem struktury, zdefiniowany jako czwórka  $G_S = (V_S, E_S, \mu_S, \xi_S)$ , gdzie:  $V_S$  – zbiór wierzchołków,  $E_S \subseteq V_S \times V_S$  – zbiór krawędzi,  $\mu_S : V_S \rightarrow L_{V_S}$  – funkcja przypisująca etykiety

wierzchołkom,  $\xi_S : V_S \rightarrow L_{E_S}$  – funkcja przypisująca etykiety krawędziom.  $L_{V_S}$  i  $L_{E_S}$  są zbiorami etykiet odpowiednio dla zbiorów wierzchołków i krawędzi. Wierzchołki takiego grafu stanowią odpowiednie elementy struktury obrazu, natomiast krawędzie określają relacje zachodzące pomiędzy tymi elementami [6].

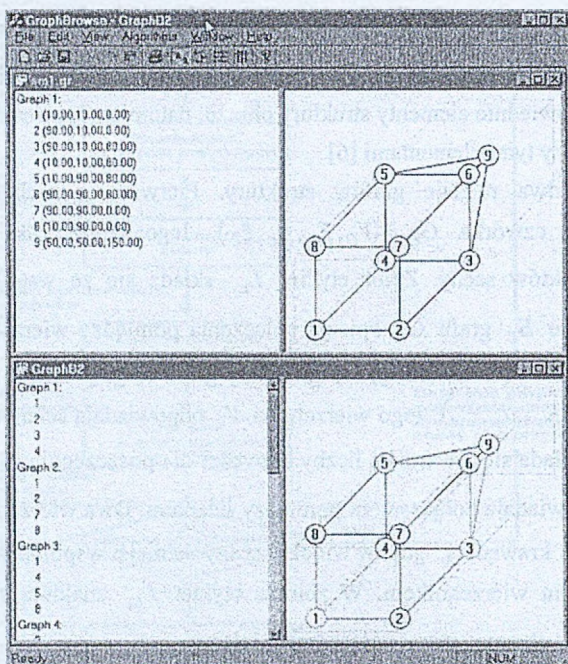
Wykorzystano dwa rodzaje grafów struktury. Pierwszy z nich to *graf konturu*, zdefiniowany jako czwórka  $G_C = (V_C, E_C, \mu_C, \xi_C)$ . Jego wierzchołki  $V_C$  odpowiadają wierzchołkom obiektów sceny. Zbiór etykiet  $L_{V_C}$  składa się ze współrzędnych  $(x, y, z)$  w scenie. Krawędzie  $E_C$  grafu  $G_C$  opisują połączenia pomiędzy wierzchołkami w scenie, czyli odpowiadają krawędziom ścian. Drugi rodzaj grafu to *graf ścian*, zdefiniowany jako czwórka  $G_F = (V_F, E_F, \mu_F, \xi_F)$ . Jego wierzchołki  $V_F$  odpowiadają ścianom obiektów sceny. Zbiór etykiet  $L_{V_F}$  składa się z wartości liczby krawędzi dla poszczególnych ścian. Krawędzie  $E_F$  grafu  $G_F$  odpowiadają połączeniom pomiędzy ścianami. Dwa wierzchołki grafu ścian są ze sobą połączone krawędzią, gdy w widoku sceny istnieje wspólna krawędź dla ścian odpowiadającym tym wierzchołkom. W zbiorze etykiet  $L_{E_F}$  znajdują się wartości kątów ściennych pomiędzy płaszczyznami połączonych ścian.

Przykładowe grafy konturu i ścian przedstawiono na rys. 6.



Rys. 6. Grafy struktury: a) graf konturu, b) graf ścian

Fig. 6. Structure graphs: a) contour graph, b) face graph



Rys. 7. Program do wizualizacji algorytmów grafowych  
 Fig. 7. Software for visualization of graph algorithms

## 7. Przetwarzanie informacji wizyjnej zapisanej w strukturach grafowych

Utworzone struktury grafowe podlegają przekształceniom prowadzącym do pozyskania struktury odpowiadającej rekonstruowanej scenie.

Istotnym krokiem jest przekształcenie grafu konturu do grafu ścian. Do tego celu służy algorytm odnajdywania wszystkich minimalnych współpłaszczyznowych cykli w grafie konturu. Algorytm opisano w [9]. Graf ścian zawiera parametry niezmiennicze względem położenia przestrzennego obiektów sceny.

Kluczowym etapem tworzenia grafów struktury sceny jest scalanie grafów struktury dla kolejnych widoków sceny. Szczegółowy opis algorytmu scalania grafów struktury zamieszczono w [9].

Rekonstrukcja sceny przebiega w  $n$  etapach, gdzie  $n$  jest liczbą punktów obserwacji. W metodzie wykorzystano algorytmy wyszukiwania subizomorfizmów w grafach [8]. W pierwszym kroku przyjęto grafy struktury dla pierwszego widoku jako początkowe grafy

struktury sceny. Proces scalania grafów struktury przebiega aż do stwierdzenia kompletności reprezentacji. Scalanie grafów struktury polega na dodaniu do grafu ścian sceny wierzchołków<sup>1</sup> pojawiających się w nowym widoku, a następnie rozszerzeniu grafu konturu sceny o wierzchołki z nowego widoku oraz przetransformowanie ich współrzędnych do pierwotnie założonego układu odniesienia sceny.

Ważną rolę w procesie rekonstrukcji odgrywa również algorytm wyznaczania następnego punktu widokowego. Za jego pomocą oszacowane jest przesunięcie ruchomej platformy do kolejnego punktu widokowego. Strategię nawigacji ruchomej platformy wokół sceny opisano w [10].

Do implementacji oraz wizualizacji algorytmów grafowych utworzono platformę programową ułatwiającą łączenie kolejnych algorytmów oraz bieżącą obserwację uzyskiwanych wyników (patrz rys. 7).

## 8. Reprezentacja wyjściowa danych

Struktury grafowe będące danymi wejściowymi, wyjściowymi, jak również wynikami cząstkowymi obliczeń przechowywane są w postaci odpowiednio zdefiniowanych zbiorów. Przyjęto dwie konwencje zapisu struktur grafowych sceny przestrzennej: konwencję konturową oraz konwencję ścienną.

<u>Zapis grafu struktury w konwencji ściennej</u>	<u>Zapis grafu struktury w konwencji konturowej</u>
<p>&lt;liczba wierzchołków w scenie (<math>n</math>)&gt;            &lt;współrzędne <math>X_1</math>&gt;            ...            &lt;współrzędne <math>X_n</math>&gt;            &lt;liczba ścian w scenie (<math>k</math>)&gt;            &lt;maksymalna liczba krawędzi pojedynczej ściany w scenie (<math>m</math>)&gt;            &lt;wierzchołki ściany <math>S_1</math>&gt;            ...            &lt;wierzchołki ściany <math>S_k</math>&gt;</p>	<p>&lt;liczba wierzchołków w scenie (<math>n</math>)&gt;            &lt;wymiar przestrzeni (3)&gt;            &lt;współrzędne <math>X_1</math>&gt;            ...            &lt;współrzędne <math>X_n</math>&gt;            &lt;krawędzie wierzchołka <math>X_1</math>&gt;            ...            &lt;krawędzie wierzchołka <math>X_n</math>&gt;</p>
<p>gdzie:            &lt;współrzędne <math>X_i</math>&gt; ::= <math>x_i, y_i, z_i, l</math>            &lt;wierzchołki ściany <math>S_j</math>&gt; ::= <math>w_{j1}, w_{j2}, \dots, w_{jk}</math><sup>2</sup></p>	<p>gdzie:            &lt;współrzędne <math>X_i</math>&gt; ::= <math>x_i, y_i, z_i</math>            &lt;krawędzie wierzchołka <math>X_i</math>&gt; ::= <math>x_{i1}, x_{i2}, \dots, x_{il}, 0</math></p>

<sup>1</sup> Wierzchołki w grafie ścian reprezentują ściany obiektów sceny

<sup>2</sup> Jeśli liczba wierzchołków ściany  $j$ , oznaczona jako  $e$ , jest mniejsza od  $k$ , to wartości  $w_{je}$  wynoszą 0.

Poniżej przedstawiono przykłady zapisu grafów struktury w obu konwencjach:

<u>Przykład konwencji ściennej</u>						<u>Przykład konwencji konturowej</u>					
12						9	3				
4.5	0	2	1			10.000000	10.000000	0.000000			
4.5	0	8.5	1			90.000000	10.000000	0.000000			
-1.3	0	11.5	1			90.000000	10.000000	80.000000			
-7	0	8.3	1			10.000000	10.000000	80.000000			
-7	0	1.7	1			10.000000	90.000000	80.000000			
-1.3	0	-1.5	1			90.000000	90.000000	80.000000			
4.5	20	2	1			90.000000	90.000000	0.000000			
4.5	20	8.5	1			10.000000	90.000000	0.000000			
-1.3	20	11.5	1			50.000000	50.000000	150.000000			
-7	20	8.3	1			2	4	0			
-7	20	1.7	1			1	3	7	0		
-1.3	20	-1.5	1			2	4	6	0		
8						1	3	5	0		
6						4	6	0			
1	2	8	7	0	0	3	5	7	0		
2	3	9	8	0	0	2	6	0			
3	4	10	9	0	0	1	5	7	0		
4	5	11	10	0	0	3	4	5	6	0	
5	6	12	11	0	0						
1	6	12	7	0	0						
7	8	9	10	11	12						
1	2	3	4	5	6						

Przytoczony przykład opisu grafu w postaci konturowej odpowiada grafowi zaprezentowanemu na rys. 7.

## LITERATURA

1. Kowalski P.: Oprogramowanie dla sterowania aktywną głowicą stereowizyjną, ZN Pol. Śl., s. Informatyka, z 38; Gliwice 2000.
2. Kowalski P.: Narzędzia programowe dla aktywnej głowicy stereowizyjnej, II Krajowa Konferencja Metody i Systemy Komputerowe w Badaniach Naukowych i Projektowaniu Inżynierskim, Kraków 1999.
3. Zonenberg D.: Protokół dla SMIS-BASIC wersja 1.0, Dokumentacja firmy SCS Design, Gliwice 1998.
4. Pojda D.: Segmentacja i pasowanie obrazów 2D dla potrzeb aktywnego stereowidzenia, Materiały II Krajowej Konferencji Metody i Systemy Komputerowe w Badaniach Naukowych i Projektowaniu Inżynierskim, Kraków 1999.
5. Philipp M.: Metoda aktywnych struktur w ustalaniu odpowiedniości obrazów stereoskopowych, Komputerowe Systemy Rozpoznawania, Wrocław 1999.
6. Deo N.: Teoria grafów i jej zastosowania w technice i informatyce, PWN, Warszawa 1980.

7. Trucco E., Verri A.: *Introductory Techniques for 3-D Computer Vision*, Prentice Hall, 1998.
8. Messmer B.T., Bunke H.: *Efficient Subgraph Isomorphism Detection: A Decomposition Approach*, *IEEE Transactions on Knowledge and Data Engineering*, vol. 12, no. 2, 2000.
9. Skabek K.: *Reconstruction of Polyhedron Objects by Structure Graphs Integration*, *Komputerowe Systemy Rozpoznawania*, Wrocław 2001.
10. Luchowski L. (ed.), Kowalski P., Tomaka A., Philipp M., Pojda D., Skabek K.: *Mobile Stereovision Strategies*, *Materiały II Krajowej Konferencji Metody i Systemy Komputerowe w Badaniach Naukowych i Projektowaniu Inżynierskim*, Kraków 1999.

Recenzent: Dr inż. Henryk Małysiak

Wpłynęło do Redakcji 31 marca 2001 r.

## Abstract

The information processing in the computer system with a mobile stereovision head was described.

As a hardware platform we used a research vehicle with a stereovision head controlled by a remote computer. The active stereovision system uses Sun Ultra 10 workstation with Sun Video Plus acquisition card (see fig. 1).

The vehicle is controlled via RS232c serial interface with a pair of radio modems. The communication is based on BNET protocol. Video signal transmission uses a wireless industrial tv link. Image acquisition is based on Sun Video Plus card under Solaris through the standard XIL library. The image resolution is 768x576 (PAL) or 640x480 (NTSC).

The methods of processing stereo raster images were presented in chapter 4.

The task of the presented system is construction of a complete 3D scene representation on the basis of stereo sequences from different points of view placed around the static scene. The process of 3D scene reconstruction is presented in fig. 5.

Structure graphs were used to represent objects of 3D scene (see chapter 6).

The graph algorithms for scene reconstruction were described in chapter 7. The process of scene reconstruction consists in searching graph matchings between the graph of the observed view and the graph of the reconstructed 3D structure. Then the integration of these structure graphs is performed for each observation point.

Finally, the notation of structure graphs is presented in chapter 8.