

AUTOMATYCZNE DOPASOWANIE SERII OBRAZÓW

Do nawigacji robota na podstawie obrazu z kamery źródłem danych jest nie jeden, lecz wiele obrazów. W opracowywanej metodzie poszukuje się obiektów powtarzających się na wielu obrazach, które przedstawiają częściowo ten sam obszar. Proponowany algorytm jest szybko i pewnie wyznacza przesunięcia obrazów na podstawie pomiarów, jeżeli spełnione są określone warunki.

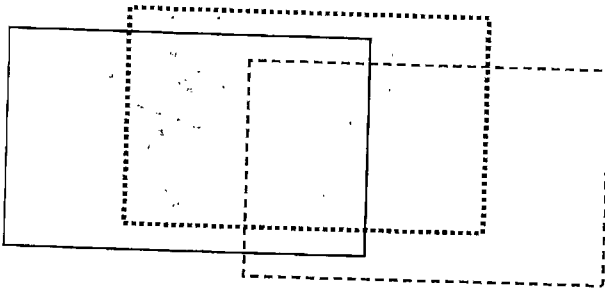
AUTOMATIC IMAGES MATCHING IN SERIE

For navigation of a mobile robot based on a camera image the base source of information are many images. In described method are processed objects, from whose each is visible in many images of partially the same view. With proposed method image displacements can be fast and reliable computed basing on measures.

1. WSTĘP

Rozpatrywanym problemem przy wyborze obrazów z wielu obrazów jest znalezienie tych samych obiektów na różnych obrazach. Etapem pośrednim tego procesu jest dopasowanie do siebie obrazów z częściowo tą samą zawartością.

Do prowadzonych prac stosowano układ obrazów obejmujący panoramę wokół robota. Schematycznie wykonane obrazów obejmujące panoramę prezentuje Rys. 1. Do znalezienia tych samych obiektów wyznacza się fragmenty obrazów, w których może wystąpić ten sam obiekt. Fragmenty te wyznacza się na podstawie znajomości przybliżonego wzajemnego przesunięcia obrazów. Celowe jest jednak wyznaczenie przemieszczenia obrazów z przyjętym dopuszczalnym błędem, w którym te same obiekty występują na różnych obrazach. Przy dopasowaniu można korzystać z wiedzy o ruchu kamery, co pozwala na zgrubne oszacowanie położenia obrazów względem siebie. W omawianym przykładzie kolejne obrazy powstały przez obrót kamery w prawo, co powoduje przesunięcie obiektów na obrazie w lewo, zniknięcie obiektów widocznych z lewej strony oraz pojawienie się nowych z prawej strony.



Rys. 1. Nakładanie kolejnych obrazów w omawianym przykładzie

2. WSTĘPNE ROZWIĄZANIE PROBLEMU

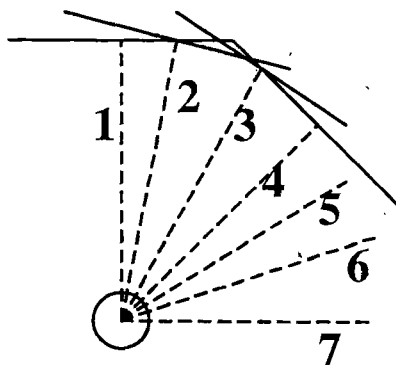
Dopasowanie obrazów można wykonać na wiele sposobów. Najprostszym wydaje się bazowanie na ruchach sterujących kamerą podczas wykonywania serii obrazów. Sposób ten nie wykorzystuje informacji zawartej w uzyskiwanych obrazach i wymaga kalibracji układu i dużej precyzji mechanizmu sterowania kamerą oraz nie może być stosowany, jeżeli robot z kamerą porusza się.

Metody wyznaczające przesunięcie na podstawie zawartości obrazów są z reguły wolne, ponieważ wyznacza się korelację obrazów lub ich fragmentów. Po wypróbowaniu wielu metod dopasowania obrazów, dość szybkie i pewne okazało się rozwiązanie przedstawione poniżej.

Znając w przybliżeniu przewidywane przesunięcie obrazu, można poszukiwać przesunięcia obiektu, który powinien być na obu obrazach, i którego wyznaczenie przesunięcia w sposób całkowicie automatyczny powinno być łatwe i pewne. Dlatego wyznacza się fragment obrazu, w którym występują wyraźne linie pionowe i poziome, przecinające się przynajmniej w jednym punkcie. Szukanie przesunięcia takiego fragmentu umożliwia znalezienie wyraźnego minimum funkcji porównującej w obu wymiarach i w tym samym punkcie. Dla każdej pary obrazów wyznacza się odpowiedni fragment biorąc pod uwagę spodziewane przesunięcie względne obrazów.

Po wyznaczeniu odpowiedniego fragmentu jednego z obrazów można spodziewać się, że na drugim obrazie wystąpi podobny, choć niekoniecznie identyczny fragment. Do znalezienia tego fragmentu stosuje się dopasowanie do szablonu przez korelację. Jeżeli w funkcji dopasowania da się znaleźć wyraźne minimum odpowiednio małej wartości, a funkcja w otoczeniu znalezionego minimum znajdują się wartości większe tylko od tego minimum, uznaje się, że współrzędne tego minimum odpowiadają przesunięciu porównywanych obrazów.

Na podstawie znajomości procesu akwizycji obrazów zakłada się, że obrazy tworzą ciąg obejmujący pełne 360° i każde jest przesunięte względem poprzedniego w prawo. Po wyznaczeniu przesunięć kolejnych obrazów wyznacza się rozmiar panoramy w pikselach, który jest parametrem charakteryzującym kamerę. Ruch kamery podczas sporządzania obrazów przedstawia Rys. 2.



Rys. 2. Schemat zmiany kamery orientacji przy wykonywaniu obrazów w pierwszych 90 stopniach

3. WYZNACZENIE ODPOWIEDNIEGO FRAGMENTU OBRAZU

Wyznaczenie fragmentu umożliwia zmniejszenie ilości obliczeń niezbędnych do wyznaczenia wzajemnego przesunięcia obrazów. Koniecznym warunkiem, jaki musi spełniać taki obszar, jest istnienie na nim co najmniej jednej ostrej linii poziomej i pionowej. Wyznaczony obszar musi mieścić się w części wspólnej obu obrazów, której granice wyznacza się na podstawie spodziewanego przesunięcia minimalnego i maksymalnego. Krawędzie poziome i pionowe wykrywa się oddzielnie maskami Prewitta wykrywającymi poziome i pionowe linie, przedstawionymi na Rys. 3.

1	0	-1
1	0	-1
1	0	-1

1	1	1
0	0	0
-1	-1	-1

Rys. 3. Maski wykrywające krawędzie pionowe i poziome

Odpowiednie miejsce charakteryzuje się istnieniem w tym samym punkcie linii poziomej i pionowej o największej sumie jasności i najmniejszej różnicy jasności. Różnica jasności linii nie może przekraczać 30 w skali 0-255. Ze względu na sposób liczenia korelacji przy zastosowaniu transformaty Fouriera, rozmiar szukanego fragmentu może być dowolny mniejszy niż okno, w którym szuka się obiektu. Wystarczająca wielkość fragmentu dla uzyskania dobrej wiarygodności dopasowania to kwadrat o rozmiarach 64 na 64 piksele.

4. WYZNACZENIE PRZESUNIĘCIA OBSZARÓW

Przesunięcie początku obszaru O1 o rozmiarach x_s i y_s względem początku większego obszaru O2 wyznacza się jako punkt, w którym istnieje minimum funkcji dwóch zmiennych:

$$D(x, y) = \sum_{i=0}^{x_s-1} \sum_{j=0}^{y_s-1} [O1(i, j) - O2(x+i, y+j)]^2 \quad (1)$$

Funkcja ta może być zapisana jako:

$$D(x, y) = \sum_{i=0}^{x_s-1} \sum_{j=0}^{y_s-1} O1(i, j)^2 - 2 \sum_{i=0}^{x_s-1} \sum_{j=0}^{y_s-1} O1(i, j)O2(x+i, y+j) + \sum_{i=0}^{x_s-1} \sum_{j=0}^{y_s-1} (O2(x+i, y+j))^2 \quad (2)$$

Pierwsza suma jest stała i jest wyznaczana wyłącznie na podstawie poszukiwanego fragmentu. Ostatnia suma jest funkcją obrazu do którego dopasowuje się fragment i można ją łatwo wyliczyć i stabilizować w celu przyspieszenia obliczeń. Środkowa suma może być zastąpiona korelacją, którą szybko wyznacza się jako iloczyn transformat Fouriera.

Korelację wyznacza się szukając dopasowania w określonych wcześniej granicach. Przy wyznaczaniu funkcji dopasowania wyznacza się okno, w którym liczy się jej wartości.

Początek tego okna zależy od wcześniej określonych przewidywanych wyników oraz od położenia charakterystycznego fragmentu.

Minimum funkcji odległości jest wykrywalne zarówno przeglądając linie pionowe, jak i poziome. Znalezienie minimum jest dobre, czyli oznacza znalezienie przesunięcia obrazów, jeżeli spełnia następujące warunki:

- pierwsza pochodna znajdowanych minimów w poziomie i w pionie jest ujemna z jednej strony wyznaczonego punktu i dodatnia z drugiej na odcinku ciągłym o długości minimum 5 pikseli;
- wartości minimów w tym ciągłym odcinku są najmniejsze dla całego porównywanego obrazu;
- współrzędne minimów wyznaczanych w liniach poziomych układają się w linię pionową, a minimów pionowych tworzą linię poziomą, dopuszcza się niewielkie odstępstwa od kierunku dokładnie poziomego lub pionowego;
- obie linie wyznaczone przez minima przecinają się w punkcie, który jest globalnym minimum;
- z wartości wyznaczonego minimum wynika średnia różnica jasności pikseli nie przekraczająca 30.

4.1. Dyskretna transformata Fouriera (DFT)

Transformatę Fouriera funkcji ciągłej wyznacza się następująco:

$$F(\omega) = \int_{-\infty}^{\infty} f(x) e^{-j\omega x} dx \quad (3)$$

Funkcję wyjściową wyznacza się z transformaty wg wzoru:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{j\omega x} d\omega \quad (4)$$

Komputery mogą przechować w tablicach wartości funkcji o dyskretnych wartościach i argumentach w ograniczonych przedziałach. Aby wyznaczyć transformatę tych funkcji, należy założyć, że funkcja jest okresowa i w tablicy znajduje się jeden pełny okres funkcji. Dyskretną transformatę Fouriera (oznaczaną jako DFT) i transformatę odwrotną funkcji dyskretniej o okresie N odwrotną wyznacza się wg wzorów:

$$F(\omega) = \frac{1}{N} \sum_{i=0}^{N-1} f(i) e^{-j\frac{\omega i}{N}} \quad (5)$$

$$f(i) = \sum_{\omega=0}^{N-1} F(\omega) e^{j\frac{\omega i}{N}} \quad (6)$$

Mimo, że ilość działań matematycznych niezbędnych do wyznaczenia transformaty jest proporcjonalna do N^2 , znane są algorytmy wykonujące się w czasie proporcjonalnym do $N \cdot \log N$, jeżeli N jest potęgą liczby 2. Algorytmy te oznaczają się jako FFT (Fast Fourier Transform) i IFFT (Inverse Fast Fourier Transform).

4.2. Wyliczenie korelacji z pomocą FFT

Transformata Fouriera obrazu jest dwuwymiarowa, ponieważ obraz należy traktować jako funkcję dwuwymiarową. Wzory transformacji dla dwu wymiarów są następujące:

$$F(u, v) = \frac{1}{M * N} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} f(i, j) e^{-j \left(\frac{ux}{N} + \frac{vy}{M} \right)} \quad (7)$$

$$f(i, j) = \sum_{u=0}^{N-1} \sum_{v=0}^{M-1} F(u, v) e^{j \left(\frac{ux}{N} + \frac{vy}{M} \right)} \quad (8)$$

Zastosowanie transformaty do wyznaczenia korelacji polega na tym, że korelacja określona wzorem:

$$D(x, y) = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} d(i, j) f(x+i, y+j) \quad (9)$$

może być wyznaczona z zastosowaniem transformaty Fouriera:

$$D(x, y) = \text{IFFT}(\text{FFT}(d(x, y)) * \text{FFT}(f(x, y)))^* \quad (10)$$

$$F(u, v) = \frac{1}{N * M} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} f(i, j) e^{-j \left(\frac{ux}{N} + \frac{vy}{M} \right)} \quad (11)$$

$$G(u, v) = \frac{1}{N * M} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} g(i, j) e^{-j \left(\frac{ux}{N} + \frac{vy}{M} \right)} \quad (12)$$

$$D(u, v) = F(u, v) * G(u, v)^* \quad (13)$$

$$D(x, y) = \text{IFFT}(D(u, v)) = \sum_{u=0}^{N-1} \sum_{v=0}^{M-1} D(u, v) e^{2\pi j \left(\frac{ux}{N} + \frac{vy}{M} \right)} \quad (14)$$

gdzie $G(u, v)^*$ jest funkcję sprzężoną do transformaty Fouriera, zaś FFT i IFFT oznaczają transformatę Fouriera i transformatę odwrotną.

Transformata Fouriera funkcji rzeczywistej jest funkcją zespoloną zmiennej rzeczywistej. Jej transformata odwrotna, transformata sumy lub iloczynu transformat funkcji rzeczywistej (np. liczona w ten sposób korelacja) jest również funkcją rzeczywistą.

Przewagą liczenia korelacji z zastosowaniem FFT nad liczeniem wprost są znane algorytmy liczenia DFT o złożoności obliczeniowej $O(N * \log N)$, gdzie N jest rozmiarem tablicy i jest potęgą liczby 2. Pozwala to na znaczne skrócenie czasu obliczeń, co jest bardzo istotne dla funkcji dwuwymiarowej określonej w względnie dużym obszarze, jaką jest obraz cyfrowy.

5. DOPASOWANIE SERII OBRAZÓW

W tym etapie wyznacza się położenie każdego z obrazów z serii w biegunowym układzie współrzędnych, którego początek znajduje się w miejscu, w którym znajduje się kamera. Obrazy muszą spełniać kilka założeń:

- każdy kolejny obraz uzyskano przy obrocie kamery w prawo, przy czym przesunięcie obrazu powinno być mniejsze niż połowa szerokości pola widzenia.
- ostatni obraz uzyskano po obrocie kamery o co najmniej 360° względem położenia przy którym uzyskano pierwszy obraz.

Wyznaczenie położenia przebiega w kilku etapach:

- wyznaczenie względnych przemieszczeń sąsiednich obrazów
- wyznaczenie rozpiętości obrazów, czyli odległości od pierwszego obrazu do ostatniego
- wyznaczenie obwodu panoramy w pikselach
- wyznaczenie położenia obrazów na podstawie przemieszczeń sąsiadów i wielkości obwodu

Taka procedura umożliwia porównanie minimalnej ilości obrazów, co pozwala na redukcję czasu obliczeń. Rozpiętość obrazów wyznacza się jako sumę przemieszczeń względnych poszczególnych obrazów.

Do wyznaczenia punktu zapętlenia wyznacza się obrazy, które obejmują częściowo obszar widoczny na pierwszym obrazie, przez wyliczenie przesunięcia pierwszego obrazu względem pozostałych. Jeżeli znana jest ilość obrazów przypadająca na obrót kamery o 360° , wystarczy porównać położenie kilku obrazów. Obwód panoramy wyznacza się jako sumę odległości od obrazu najbliższego pierwszemu z lewej strony, a więc ostatniego przed pełnym obrotem i sumy odległości od pierwszego obrazu do ostatniego przed pełnym obrotem kamery.

6. EKSPERYMENTY

Dopasowanie jest możliwe poza wyznaczonym zakresem przy nieznacznym wydłużeniu czasu obliczeń. Krytyczne pod względem czasowym jest wyznaczenie korelacji dwu obrazów. Dla uproszczenia obliczeń i uniknięcia powtarzania wyznaczania korelacji w większym obszarze, wylicza się ją w prostokącie 512×256 pikseli, co obejmuje rozsądny przedział przesunięć obrazów. Samą korelację wyznacza się jako iloczyn dyskretnych transformat Fouriera. Zastosowanie FFT powoduje niezależność prędkości obliczeń od wielkości przedziałów mniejszych niż wyznaczone okno, oraz wielokrotne zmniejszenie czasu obliczeń. Na komputerze z procesorem Pentium IV 1400 MHz dopasowanie obrazów z zastosowaniem DFT w oknie 512×256 pikseli trwa 4 sekundy, zaś dopasowanie bez DFT, w oknie 30 na 20 pikseli – ok. 5 sekund. Dopasowanie obrazów w oknie 512×256 pikseli bez stosowania FFT trwa kilka godzin.

Przy zbyt wąskich przedziałach przewidywanych współrzędnych dla rzeczywistych obrazów okazywało się, że nie zawsze przesunięcie mieściło się w wyznaczonych granicach.

Eksperymenty pokazały, że opisana metoda może działać stabilnie dla obrazów testowych płaskich, trójwymiarowych pochodzących z symulacji oraz rzeczywistych. Dobre wyniki uzyskano, jeżeli:

możliwe jest dopasowanie obrazów przesuniętych względem siebie więcej niż oczekiwane przesunięcie;

możliwe jest znalezienie charakterystycznego fragmentu w części obrazu, która jest wspólna przy przesunięciach mieszczących się tylko częściowo w przewidywanym zakresie.

Po optymalizacji i zastosowaniu FFT dopasowanie serii (32-37 obrazów) zajmuje na komputerze z procesorem Pentium IV z zegarem 1.4 GHz 4-5 minut. W zamieszczonych tabelach znajdują się wyniki uzyskane dla:

obrazów z symulacji o rozdzielczości 640x480 (Tabela 1);

obrazów pochodzących z wnętrza pomieszczenia (Tabela 2).

Ilość obrazów	Wielkość panoramy
37	2116
37	3141

Tabela 1. Wyniki dla danych symulowanych 3D

Ilość obrazów	Wielkość panoramy	Średnia odległość	Średnie odchylenie odległości	Błąd
33	3737	123	4,6	3,7%
32	3736	120	2,6	2,2%
32	3686	119	2,5	2,1%

Tabela 2. Wyniki dla danych rzeczywistych uzyskanych przy równych obrotach kamery

7. WNIOSKI

Stosując znane techniki dopasowania fragmentu obrazu do szablonu udało się opracować metodę wyznaczania obwodu panoramy wokół kamery z pojedynczych obrazów. Opracowana metoda charakteryzuje się małym błędem względnym i krótkim czasem działania. Dla obrazów uzyskanych przez obrót kamery o ten sam kąt średni błąd przesunięcia nie przekracza 4%.

LITERATURA

- [1] W. K. Pratt: *Digital image processing*; John Willey & sons inc., New York 2001.
- [2] Ch. D. Watkins, A. Sadun, S. Marenka: *Nowoczesne metody przetwarzania obrazu*; Wydawnictwa Naukowo-Techniczne, Warszawa 1995.