

8. AUTOMATYCZNA SEGMENTACJA KOLOROWYCH SEKWENCJI WIZYJNYCH

Niniejsze ćwiczenie ma na celu zapoznanie się z segmentacją sekwencji wizyjnych opartych na zasadzie rozrostu.

W ćwiczeniu wykorzystywana jest aplikacja **segmentation** umożliwiająca przeprowadzenie procesu segmentacji sekwencji wizyjnej na zasadzie rozrostu.

Wprowadzenie teoretyczne (źródło - praca *automatyczna segmentacja kolorowych sekwencji wizyjnych* autor Tomasz Bartczak, opiekun dr inż. Maciej Bartkowiak):

1. Zasada działania algorytmu rozrostu

Poniższy opis przedstawia implementację automatycznej segmentacji kolorowych sekwencji wizyjnych. Podstawowym problemem w pełni automatycznej segmentacji jest identyfikacja regionów w pierwszej ramce sekwencji, a następnie śledzenie ich w kolejnych ramkach. Algorytm przedstawiony na rysunku można podzielić na dwie podstawowe części: rozrost regionu przy użyciu kryterium jednorodności kolorów, a następnie śledzenie obiektów z wykorzystaniem kryterium ruchu oraz porównania ramek.

Po wstępnym przetworzeniu obrazu rozpoczyna się statyczna segmentacja pierwszej ramki sekwencji. W ten sposób uzyskuje się regiony z jednoczesną ich indeksacją, które są następnie poszukiwane w kolejnej ramce. Umożliwia to rozrost nie tylko w dwóch wymiarach płaszczyzny obrazu, ale również w trzecim wymiarze, jakim stają się kolejne ramki sekwencji. Tak przetworzony obraz jest oczyszczany z nieprawidłowych regionów, to znaczy z takich, których czas trwania wynosi jedną ramkę lub są regionami bardzo małymi przez cały czas ich trwania.

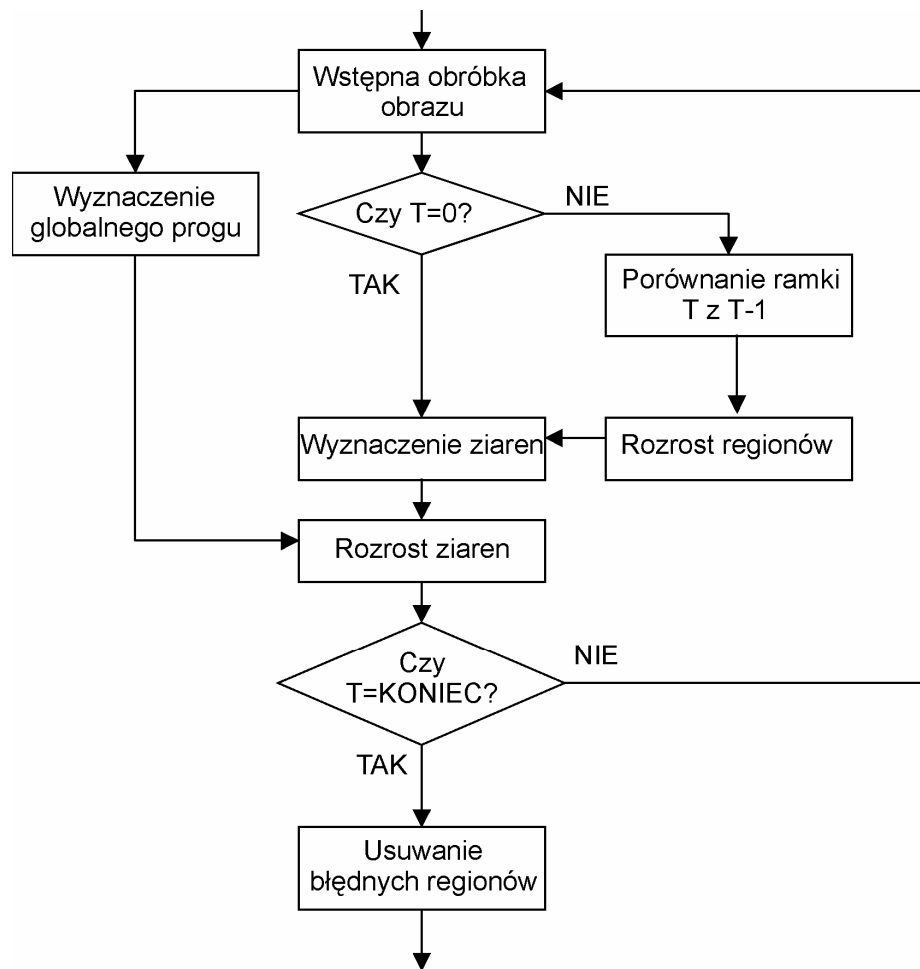


Diagram algorytmu automatycznej segmentacji sekwencji wizyjnej

2. Filtracja obrazów wejściowych

Zanim rozpocznie się segmentacje obraz należy przefiltrować w celu usunięcia szumów. Zakłócenia w obrazie mogą spowodować powstanie bardzo małych i szybko zmieniających się punktów, które przez program mogą zostać zidentyfikowane jako obiekty, a tego należy uniknąć. Do złagodzenia wpływu szumu można użyć filtru medianowego o masce 5x5:

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

dla każdej ze składowych przestrzeni $Y \ C_R \ C_B$, który odrzuca skrajne wartości, a za wynik filtracji przyjmuje dokładnie środkową z wartości w ciągu rosnącym obejmowanym przez maskę. Zatem mając punkt $s(x, y)$ oraz maskę $k \times k$:

$$S = \{s_1, s_2, \dots, s_{2k-1}\}$$

gdzie S jest zbiorem punktów s_i jednej ze składowych przestrzeni kolorów $Y \ C_R \ C_B$ ograniczonych przez maskę oraz zakładając, iż:

$$s_1 \leq s_2 \leq s_3 \leq \dots \leq s_{2k-2} \leq s_{2k-1}$$

jest ciągiem rosnącym, to mediana tego zbioru określona jest jako:

$$\text{med } S = s_k$$

Pojawia się również problem ze zdecydowaniem składowej chrominancji C_R oraz C_B w stosunku do luminancji. Format, jaki pojawia się na wejściu jest związany z modelem próbkowania chrominancji 4: 2: 0, oznacza to, iż na cztery wartości luminancji przypada po jednej wartości chrominancji C_R oraz C_B . Problem został zmniejszony przez zastosowanie prostej interpolacji dla składowej chrominancji postaci:

$$\frac{1}{32}[-1 \ 0 \ 9 \ 16 \ 9 \ 0 \ -1]$$

3. Określenie wartości progu akceptacji punktów podczas rozrostu

Dla wcześniej przefiltrowanego obrazu wyznaczony zostaje globalny próg akceptacji kolejnych punktów podczas rozrostu ε_g . Przyjmijmy, iż V jest zbiorem

wszystkich punktów w sekwencji oraz, że I jest równoważne z jedną ze składowych przestrzeni kolorów $Y \subset \mathbb{R}^3$. Zatem globalna średnia wartość I wynosi:

$$\eta = \frac{1}{M} \sum_{x,y,t \in V} I(x,y,t) \quad (3.5),$$

gdzie M jest całkowitą liczbą punktów. Globalna wariancja σ^2 wyraża się wzorem:

$$\sigma^2 = \frac{1}{M} \sum_{x,y,t \in V} (I(x,y,t) - \eta)^2 \quad (3.6).$$

Dynamiczny zakres μ dla wszystkich punktów wynosi:

$$\mu = \max_{x,y,t \in V} I(x,y,t) - \min_{x,y,t \in V} I(x,y,t) \quad (3.7).$$

W literaturze proponowano by globalny próg przydzielany był na podstawie skalowania dynamicznego zakresu:

$$\varepsilon_g = k \frac{\mu}{\sigma + 1} \quad (3.8),$$

gdzie $k > 1$ stanowi parametr czułości. Z obserwacji, jaką przeprowadzono w [3] wynika, iż dobre rezultaty uzyskuje się dla $k \approx 2$.

Na rysunkach 3.2 oraz 3.4 przedstawione zostały pierwsze ramki sekwencji odpowiednio `claire.tif` i `salesman.tif` poddane procesowi segmentacji z różnymi wartościami parametru czułości k (1.1, 1.6, 2.1, 2.6 oraz 3.1), natomiast na rysunkach 3.3 oraz 3.5 przedstawiono granice regionów. Rysunki 3.6 oraz 3.7 przedstawiają wykresy zależności liczby regionów w ramce od parametru czułości. Segmentacja

przeprowadzona została dla 20 ramek. W przypadku segmentacji sekwencji *claire.cif* (rysunek 3.2, 3.3 oraz 3.6), wraz ze wzrostem parametru k zmniejszała się znacząco liczba regionów jednocześnie nie wprowadzając znaczących błędów w regionach. Dlatego w przypadku tej sekwencji można zastosować współczynnik k o wartości ponad 3. Inaczej wygląda to dla sekwencji *salesman.cif* (rysunek 3.4, 3.5 oraz 3.7), gdzie znaczące błędy pojawiają się dla k większego od 2.1. Widoczne jest to bardzo dobrze na rysunku 3.4d oraz na granicach jego regionów przedstawionych na rysunku 3.5d. Regiony odpowiedzialne za podział na włosy i półkę z książkami częściowo połączyły się. Wynika to z faktu niewystarczającej różnicy wektorów kolorów obu regionów w stosunku do wartości progu akceptacji punktów podczas rozrostu. Podobne problemy pojawiają się pomiędzy lewą dłonią a koszulą w obrazie 3.4d oraz 3.4e, jak również pomiędzy pudełkiem trzymanym w prawej dłoni a tłem w rysunku 3.4e.

Optymalny dobór współczynnika k ze wzoru (3.8) należy uzależnić od dwóch parametrów: liczby regionów w ramce, która powinna być jak najmniejsza oraz dokładności regionów w stosunku do rzeczywistych obiektów, dlatego dobrym rozwiązaniem jest ustawienie parametru czułości na wartość bliską 2, która sprawia, iż regiony dobrze odzwierciedlają rzeczywiste obiekty.

W ten sposób określone zostają globalne progi, do jakich ziarno może się rozrastać dla wszystkich składowych przestrzeni barw $Y C_R C_B$. Zatem region K jest rozszerzony o punkt P , jeżeli P jest sąsiadem regionu K oraz norma Euklidesowa różnicy kolorów ziarna i punktu P jest mniejsza od normy Euklidesowej z progów dla wszystkich składowych $Y C_R C_B$:

$$\sqrt{\left((Y_Z - Y_P)^2 + (C_{RZ} - C_{RP})^2 + (C_{BZ} - C_{BP})^2\right)} \leq \sqrt{\left(\varepsilon_{gY}^2 + \varepsilon_{gC_R}^2 + \varepsilon_{gC_B}^2\right)} \quad (3.9),$$

gdzie ε_{gY} , ε_{gC_R} i ε_{gC_B} są odpowiednio progami akceptacji dla luminancji oraz chrominancji typu R oraz B, $Y_Z C_{RZ} C_{BZ}$ są składowymi przestrzeni barw dla ziarna, natomiast $Y_P C_{RP} C_{BP}$ są składowymi przestrzeni barw dla punktu sąsiadującego z ziarnem.

4. Wyznaczanie ziaren

Należy założyć, iż obraz składa się z wielu regionów, które posiadają ziarno, czyli punkt, od którego można rozpocząć rozrost, w celu stworzenia regionu. Segmentację rozpoczynamy od wyznaczenia tychże ziaren, które będą zaczątkami możliwych regionów. Dokonać tego można obliczając wartość gradientu dla każdego punktu obrazu, przy użyciu operatora Sobel'a [2], estymowane jest to przy użyciu masek:

$$I_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad \text{oraz} \quad I_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (3.10),$$

gdzie I_x i I_y są odpowiednio pochodnymi cząstkowymi dla współrzędnej poziomej oraz pionowej.

Gradient wyznaczany jest dla każdej składowej przestrzeni barw Y C_R C_B . Wymiar gradientu opisuje równanie:

$$|\nabla I(x, y)| = \sqrt{(I_x^2 + I_y^2)} \quad (3.11).$$

Następnie należy zastąpić gradient dla poszczególnych składowych przestrzeni kolorów, w postaci wartości gradientu luminancji i chrominancji jedną wartością. Zastosować można w tym celu zastąpienie składowych ich normą Euklidesową:

$$G(x, y) = \sqrt{(\nabla Y(x, y))^2 + \nabla C_R(x, y)^2 + \nabla C_B(x, y)^2} \quad (3.12)$$

Tak wyznaczone ziarna tworzą maskę o wymiarach obrazka, w której wartość 1 przypisywana jest punktom, które można uznać za ziarna, natomiast 0 określać będzie punkty, które nie mogą stać się ziarnami.

5. Rozrost ziaren

Rozrost ziarna dokonuje się dla punktu $P_i(x,y)$ o minimalnej wartości gradientu określonego we wzorze (3.12), który jak dotąd nie został zaliczony do jakiegokolwiek z regionów. Ziarno P_i , oznaczone niepowtarzalną etykietą W_i , tworzy samodzielny region, który rozszerza się o punkty sąsiednie jak dotąd nie należące do któregoś z regionów, gdy spełniony został warunek (3.9). Rozrost najlepiej dokonywać jedynie w kierunku poziomym lub pionowym, czyli dla punktów $P(x+1,y)$, $P(x-1,y)$, $P(x,y+1)$, $P(x,y-1)$, w ten sposób uniknie się stworzenia regionu, który połączony będzie jedynie rogami dwóch punktów, ułatwia to implementacje algorytmów, które wykorzystują przeszukiwanie regionów.

Połączone w region punkty oznaczone zostają etykietą ziarna i zyskują część jego właściwości, to znaczy, iż można dokonać rozrostu również i od dołączonych do regionu punktów, gdy spełniony zostanie warunek (3.9). Jednak warunek ten musi zostać spełniony w stosunku do składowych przestrzeni barw ziarna i punktu dołączanego, a nie składowych punktu, który został wcześniej dołączony do regionu i w ten sposób zyskał część właściwości ziarna.

Punkty, które po rozroście ziaren nie zostały przydzielone do regionów i posiadają przynajmniej trzech sąsiadów zaliczonych do jakiegoś regionu, można porównać do sąsiadujących z nimi regionów, następnie wybierając najbardziej zbliżony region, czyli ten o najmniejszej normie Euklidesowej z różnicy ziarna regionu i punktu, należy połączyć punkt z wybranym regionem:

$$\min_{i \in R} \sqrt{(Y_Z - Y_i)^2 + (C_{RZ} - C_{Ri})^2 + (C_{BZ} - C_{Bi})^2} \quad (3.13),$$

gdzie Y_Z C_{RZ} C_{BZ} są składowymi przestrzeni barw dla ziarna, do którego przyłączony jest punktu $P(x,y)$, od którego możliwy jest rozrost, natomiast Y_i C_{Ri} C_{Bi} są składowymi przestrzeni barw dla punktu $P'(i)$ kandydującego do regionu i jednocześnie sąsiadującego z $P(x,y)$, gdzie $i = (x+1,y)$, $(x-1,y)$, $(x,y+1)$ oraz $(x,y-1)$.

Pozostałe punkty są przyłączane do regionów zgodnie z podpunktem 3.5 przy zwiększonym progu akceptacji punktów podczas rozrostu o 5%. Wartość progu zwiększana jest do czasu przydzielenia wszystkich punktów do regionów.

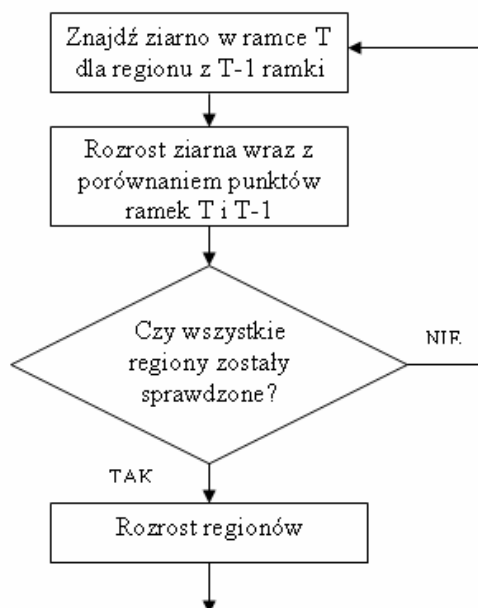
W wyniku tak działającego algorytmu powstają regiony dla pierwszej ramki sekwencji.

Wynik działania segmentacji dodatkowo poprawiony jest przez uśrednienie poszczególnych wartości składowych przestrzeni barw dla każdego z regionów. Dla każdej ze składowych przeprowadzana jest estymacja wartości średniej zgodna z równaniem:

$$\overline{I}_{W_i} = \frac{1}{N_{W_i}} \sum_{x,y \in W_i} I(x,y) \quad (3.14),$$

gdzie W_i jest etykietą regionu, N_{W_i} sumą punktów w regionie W_i , natomiast I jest jedną ze składowych Y C_R C_B .

6. Śledzenie regionów



Rys.3.6. Diagram algorytmu śledzenia regionów

Segmentacja drugiej ramki, jak również i kolejnych, rozpoczyna się od porównania dwóch ramek. Dla każdego regionu W_i z ramki wcześniejszej (T-1) poszukujemy się adekwatnego ziarna w ramce T. Poszukiwania ziarna ograniczone są do obszaru, jaki reprezentuje region W_i , tworząc maskę dla ramki T. Punkt, który osiągnie minimum zgodne ze wzorem (3.13) i jednocześnie spełniona jest nierówność:

$$\sqrt{\left((Y_Z - Y_P)^2 + (C_{RZ} - C_{RP})^2 + (C_{BZ} - C_{BP})^2\right)} \leq \frac{1}{2} \sqrt{\left(\varepsilon_{gY}^2 + \varepsilon_{gC_R}^2 + \varepsilon_{gC_B}^2\right)} \quad (3.15),$$

zostaje uznany za ziarno W_i dla ramki T, będącego kontynuacją regionu W_i z ramki T-1.

Rozrost tak uzyskanego ziarna odbywa się podobnie jak opisano to w podpunkcie 3.5. Zmianie uległ jedynie obszar rozrostu ograniczony do wcześniej omówionej maski regionu oraz próg akceptacji punktu do regionu, zgodny z równaniem (3.15). Połowa wartości normy euklidesowej dopuszczalnych odległości, wynika z faktu, iż kolor regionu z ramki T-1 stanowi wartość średnią z całego regionu. Zatem ziarno mogłoby być w tym regionie jedną ze skrajnych wartości. By uniknąć wyszukania ziarna w ramce T spoza dopuszczalnej granicy regionu z ramki T-1, należy przyjąć próg o połowę mniejszy. W ten sposób mamy pewność, iż ziarno znalezione w ramce T spełnia parametry rozrostu dla regionu z ramki T-1.

Po zakończeniu poszukiwań wszystkich regionów z ramki T-1 w ramce T, znalezione regiony ponownie się rozrastają, lecz tym razem całkowicie zgodnie z podpunktem 3.5. Następnie należy usunąć pojedyncze punkty wewnątrz regionów zgodnie z zależnością (3.13).

Dobre rezultaty uzyskuje się ponownie rozszerzając regiony o punkty dotychczas nie zaliczone do jakiegokolwiek z regionów, lecz tym razem zwiększając próg o 10%. Pozostałe punkty nie zrzeszone z żadną etykietą, tworzą regiony, które dotychczas nie występowały w obrazie, np. powstałe w wyniku znaczących zmian światła lub cienia jak również całkowicie nowe obiekty. Dlatego dla tych regionów stosuje się operacje opisane w podpunkcie 3.4, a następnie te zawarte w podpunkcie 3.5.

Po zakończeniu segmentacji, tak jak to ma miejsce przy pierwszej ramce, każdy region ma uśredniane wartości składowych przestrzeni barw zgodnie z równaniem (3.14).

Obsługa programu:

W ćwiczeniu wykorzystywana jest implementacja algorytmu segmentacji sekwencji wizyjnej poprzez rozrost (program **segmentation** autorstwa Pana Tomasza Bartzaka). Program jest typowym programem konsolowym.

Dane wejściowe potrzebne do segmentacji to:

- a) Format pliku wejściowego (`-xCIF`), zdefiniowane są tylko trzy możliwe typy:
QCIF, CIF oraz 4CIF.
- b) Model próbkowania chrominancji (`-4xx`),
- c) Nazwa pliku wejściowego (`-i nazwa_pliku`).

Prawidłowe podanie tych trzech parametrów po nazwie `Segmentation` jest wystarczające do przeprowadzenia segmentacji. Zdefiniowane zostały również dodatkowe opcje:

- a) Możliwość rozpoczęcia segmentacji od ramki numer N (`-s N`).
- b) Segmentacja wyznaczonej liczby ramek większej od 1 (`-l liczba_ramek`).
- c) Określenie przez użytkownika nazwy pliku wyjściowego (`-o nazwa`), standardowo jest to `Seg`.
- d) Określenie pozycji regionu (`-a x` oraz `-b y`) wybranego do prezentacji. Do pliku `Seg_isolate` zapisywana jest sekwencja z wyszczególnionym regionem, który zawiera punkt $(-a, -b)$.
- e) Zdefiniowanie przez użytkownika wartości progu akceptacji punktów podczas rozrostu, standardowo odbywa się to automatycznie zgodnie z opisem.
- f) Pomoc dostępna jest przy użyciu opcji `-h`.

Polecenia:

Uwaga! W roku 2009 realizowany jest jedynie punkt 1 i 2 poleceń!

1. Automatyczna segmentacja sekwencji wizyjnej

Proszę dokonać segmentacji sekwencji wizyjnej clair.cif oraz salesman.cif dla kilku wybranych wartości globalnego progu. Proszę się odnieść do powstałych błędów:

- a) błędy powstałe w regionach o zbliżonych kolorach;
- b) błędy związane z gubieniem wąskich bądź bardzo małych obiektów;
- c) prawidłowe „trzymanie” nieruchomych regionów.

2. Odpowiedzieć i opisać w niezależnym dokumencie dwa poniższe tematy:

- a) **Assessment of the state-of-the-art in image and video segmentation**
- b) **Are there publicly available datasets containing ground truth masks for videos?**

Uwaga! Bardzo ważne – Punkt 2 jest równocześnie pracą zaliczającą przedmiot Biometria i nadzór wizyjny! Proszę oddać ten dokument w wersji elektronicznej pocztą mailową do końca semestru! Ostatecznym dniem przyjmowania dokumentów jest ostatni dzień aktualnego semestru! Po tym terminie dla osób, które nie oddadzą tego dokumentu będzie przeprowadzony egzamin w tradycyjnym trybie tj. część pisemna egzaminacyjna (odpowiedzi na pytania) oraz część ustna. W dokumencie oceniane będą następujące czynniki: jakość informacji zawartych w dokumencie, aktualność informacji, referencje, styl i uporządkowanie dokumentu. Proszę zwrócić uwagę na redakcję tego dokumentu!

3. Implementacja prostego algorytmu segmentacji sekwencji wizyjnej w Matlabie

- a) Dokonać implementacji proponowanego algorytmu automatycznej segmentacji sekwencji wizyjnej w programie Matlab.
- b) Przeprowadzić analizę statystyczną segmentacji regionów z wykorzystaniem następujących parametrów:

Wartość średnia liczby regionów na ramkę, określona jest wzorem:

$$\overline{L}_T = \frac{1}{N_{wi}} \sum W_i(T)$$

gdzie N_{W_i} jest liczbą ramek w sekwencji, natomiast W_i jest regionem w ramce T , gdzie $T \in (1; N_{W_i} - 2)$. Nie można uwzględniać pierwszej i ostatniej ramki, które zawierają odpowiednio tylko nowe i stare regiony, w ten sposób mogłyby wprowadzać drobne zniekształcenie wyników, dotyczy to wszystkich obliczeń statystycznych.

Rozstęp określony jest równaniem:

$$R = L_{MAX} - L_{MIN}$$

gdzie L_{MAX} oraz L_{MIN} są odpowiednio największą oraz najmniejszą liczbą regionów w ramce.

Wariancja wyrażona jest wzorem:

$$\sigma^2 = \frac{1}{N_{W_i}} \sum (W_i(T) - \bar{L})^2$$

gdzie N_{W_i} jest liczbą ramek w sekwencji, natomiast $W_i(T)$ jest regionem w ramce T , gdzie $T \in (1; N_{W_i} - 2)$. \bar{L} jest wartością średnią opisaną w równaniu (5.1). Zatem przeciętne odchylenie od średniej określone jest równaniem:

$$u = \frac{1}{N_{W_i}} \sum |W_i(T) - \bar{L}| = \frac{1}{\sqrt{N_{W_i}}} \sigma$$

Obliczenia statystyczne wykonać dla sekwencji salesman.cif oraz claire.cif w trzech wariantach:

- dla wszystkich regionów w ramce;
- dla nowopowstałych regionów w ramce;
- oraz dla tych regionów, które są kontynuacją z poprzedniej ramki.

Uwagi organizacyjne:

1. Katalogiem roboczym do umieszczania własnych plików jest **x:**. Przed przystąpieniem do ćwiczeń należy usunąć wszystkie pliki z tego katalogu.
-

