ІНФОРМАТИКА, ОБЧИСЛЮВАЛЬНА ТЕХНІКА ТА АВТОМАТИЗАЦІЯ

УДК 658.5.011.56:004.932.001.57

С.Г. Антощук, д.т.н., проф.,
О.Ю. Бабілунга, к.т.н., доц.,
А.О. Ніколенко, к.т.н., доц.,
О.В. Ткаченко, асист.
Одеський національний політехнічний університет

АДАПТИВНА ЛОКАЛІЗАЦІЯ СИМВОЛЬНИХ НАПИСІВ НА ЗОБРАЖЕННЯХ МЕТОДОМ ВЕЙВЛЕТ-АНАЛІЗУ

Запропоновано підхід до адаптивної локалізації символьної інформації на зображеннях шляхом аналізу функції енергії сигналу в просторі вейвлет-перетворення, що дозволило розробити модель формування образів об'єктів на зображенні та відповідні інформаційні технології.

Актуальність та постановка проблеми. Серед завдань, що вирішуються в області комп'ютерного зору, найбільш класичним і актуальним є завдання розпізнавання символьної інформації (СІ), яке виникає при побудові автоматичних систем обробки візуальної інформації, що використовуються для автоматизації документообігу, обліку транспорту за номерами державної реєстрації, пошуку символьної інформації в інтернет-колекціях зображень тощо. В деяких застосуваннях (наприклад системи оптичного розпізнавання тексту) інтерпретація зображень з СІ не викликає труднощів, однак у ряді додатків існують проблеми, пов'язані з цільовим призначенням кожної з комп'ютерних систем розпізнавання СІ (КСР СІ). При розпізнаванні СІ необхідно враховувати різноманітність положення окремих символьних областей і фонових заповнень, а також форм і розмірів символів, їх можливі спотворення та ін. Вирішення цих проблем можливе шляхом використання адаптивних технологій, що враховують структуру зображення, яке містить СІ. Для цього необхідна модель формування образів об'єктів, параметри якої істотно впливають на ефективність роботи КСР СІ і її складових частин. Створення моделі формування образів об'єктів на зображенні дозволить врахувати інформацію на різних рівнях представлення об'єкта, його структурні й геометричні особливості.

Розпізнавання СІ як правило виконується в два етапи: локалізація СІ і власне її розпізнавання. При цьому саме якісна локалізація (особливо при складному фоні) визначає образи структури зображення в цілому і окремих символів. Це дозволяє перейти до моделей, побудованих на знаннях, розглядати зображення за допомогою відношень «частина–ціле», забезпечує можливість адаптації до різних застосувань, скорочує час інтерпретації зображень за рахунок проведення цілеспрямованого пошуку СІ при високій надійності її розпізнавання.

Аналіз існуючих рішень щодо адаптивної локалізації СІ. Локалізація може розглядатися як різновид сегментації, яка полягає в представленні оброблюваного зображення у вигляді сукупності областей, що задовольняють заданий критерій однорідності, тому були проаналізовані два основні підходи до сегментації текстових областей. Перший – «зверху-вниз», при якому документ рекурсивно ділиться на дрібні частини, реалізується такими алгоритмами сегментації [1]: проекцій, гістограмного аналізу, алгоритмами, що побудовані на правилах і на просторових перетвореннях (Фур'є, Хо). Для цих алгоритмів характерна висока швидкодія і перешкодостійкість, проте ефективність безпосередньо залежить від апріорних знань про структуру і клас документа. Другий підхід – «знизу-уверх» полягає в потоншенні елементів, об'єднанні їх в зв'язкові області й потім – у смислові структури. Найбільш популярні алгоритми цієї групи: математичної морфології; нарощування областей; на основі розрахунку градієнта. Алгоритми цієї групи не вимагають знаннь про початкову структуру документа, проте мають низьку перешкодостійкість.

Метою даної роботи є розробка методу адаптивної локалізації символьної інформації на зображеннях з використанням вейвлет-перетворення на основі моделі образів СІ на зображенні, яка враховує структурні та масштабні особливості представлення.

Викладення основного матеріалу. Формування моделі СІ. При побудові адаптивних алгоритмів обробки зображень з СІ необхідно при виконанні кожного етапу обробки враховувати: початкове представлення (форму вхідного зображення), алгоритми обробки (їх параметри) і вихідне представлення (згідно з метою етапу). Вихідне представлення повинне відповідати моделі формування образів об'єктів, адекватно відображати початкове представлення і містити значущу інформацію (топологічні та геометричні ознаки) про об'єкт інтересу, яка визначається задачею, що розв'язується. Зображення з СІ може бути розглянуто як ієрархічно складно організований об'єкт, на якому можна виділити деякі об'єкти та області з символьною (текстовою) інформацією (рис. 1). Таке представлення при вирішенні

© С.Г. Антощук, О.Ю. Бабілунга, А.О. Ніколенко, О.В. Ткаченко, 2008

завдання аналізу інформації згідно зі стандартом MPEG-7 передбачає декомпозицію зображення на складові частини, тобто виділення областей, рядків у них, окремих слів і символів [2, 3].

Рис. 1. Приклад зображення, що містить символьну інформацію

Фрагмент зображення, що містить СІ, при аналізі можна розглядати як набір стаціонарних областей, що характеризуються відповідними сегментними ознаками, характерними для опису СІ. Рядок такої області з СІ (на рис. 1 показано стрілкою) може бути промодельований за допомогою послідовності імпульсів з квазіпостійними періодом та тривалістю. У такому разі математична модель рядка з окремим символом:

$$u(x) = 1(x - x_0) - 1(x - x_0 - d),$$

де x_0 – координата початку символу;

d – тривалість символу.

Тоді, якщо напис містить N символів, модель рядка має вигляд:

$$U(x) = \sum_{i=1}^{i=N} \left(1(x - x_0 - (i-1)T) - 1(x - x_0 - (i-1)T - d)\right), \quad (1)$$

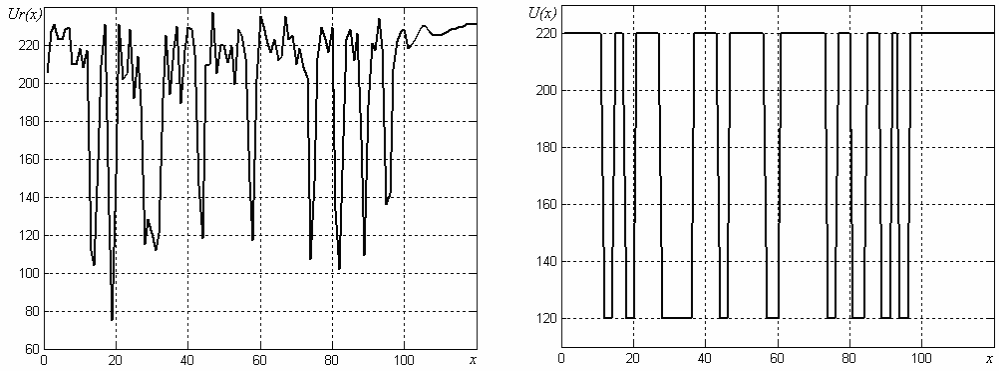
де T – період розташування символів.

Якщо урахувати випадкові зміни періоду (ε_{Tk}) та тривалості (δ_i), які визначаються особливостями шрифту напису, то модель (1) приймає вигляд:

$$U(x) = \sum_{i=1}^{i=N} \left(1(x - x_0 - (i-1)T - \sum_{k=2}^{i} \varepsilon_{Tk}) - 1(x - x_0 - (i-1)T - \sum_{k=2}^{i} \varepsilon_{Tk} - d - \delta_i)\right). \quad (2)$$

На рис. 2 представлено фрагмент рядка реального зображення (рис. 2, *а*) та графічне зображення моделі (2) (рис. 2, *б*), яка має подібні геометричні характеристики. Наведена модель рядка зображення враховує ієрархічну структуру СІ, яка представлена у вигляді сукупності фон–підоб'єкт 1-го рівня ієрархії (напис)–підоб'єкти 2-го рівня ієрархії (символи) [3]. Запропонована модель у подальшому використовувалась для проведення аналітичного дослідження та імітаційного моделювання процесу локалізації СІ.

Аналіз просторово-частотних властивостей моделі СІ. Для забезпечення адаптивної локалізації СІ пропонується враховувати просторово-частотні властивості, які оцінюються за допомогою математичного апарата вейвлет-аналізу. Для вирішення поставленого завдання вейвлет-функції повинні забезпечувати підкреслення перепадів інтенсивності окремих символів (підоб'єктів) з високою перешкодостійкістю та можливість регулювання детальності об'єктів за рахунок частотно-селективних властивостей. Аналіз різних базисів вейвлет-перетворення (ВП) показав, що такі вимоги задовольняють дійсні вейвлети, які є непарними симетричними функціями, мають компактний або ефективний носій. В основу такого адаптивного базису покладено «розщеплену» функцію Гауса [4].



а) *б)*

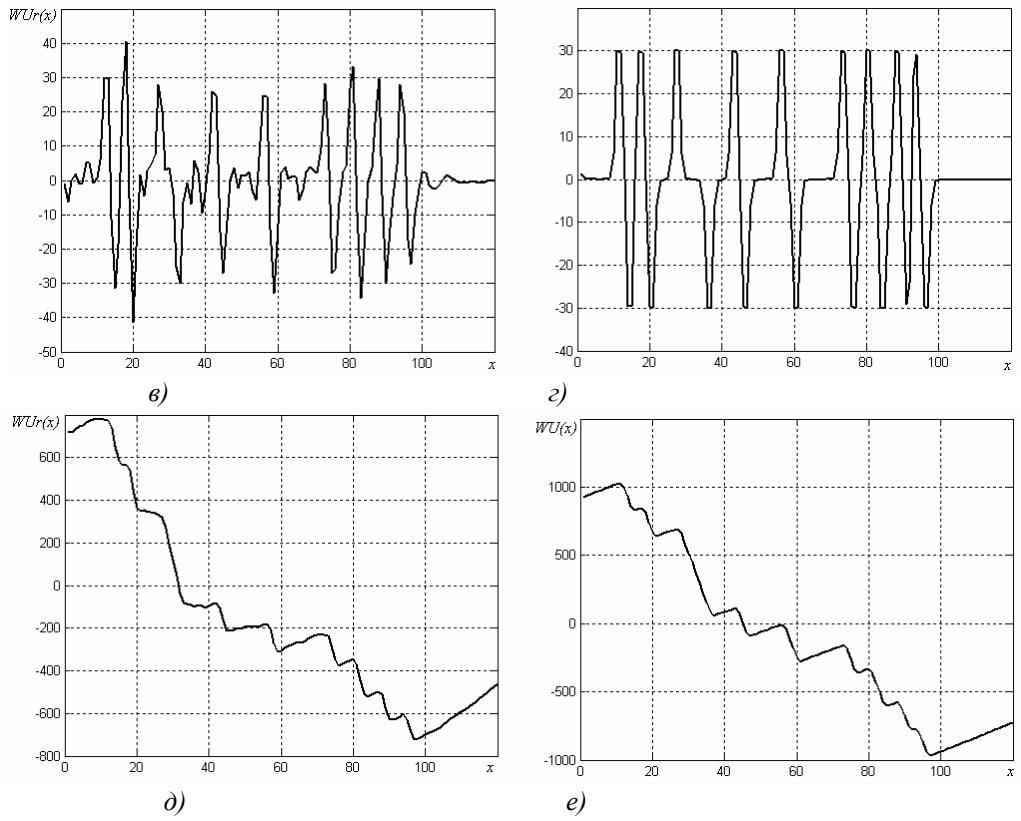


Рис. 2. Фрагмент рядка з СІ реального зображення (а); модель рядка з СІ (б); ВПРГ з масштабом s1 рядка з СІ реального зображення (в); ВПРГ з масштабом s1 моделі рядка з СІ (г); ВПРГ з масштабом s2 рядка з СІ реального зображення (д); ВПРГ з масштабом s2 моделі рядка з СІ (е), (s1<s2)

У цьому випадку базисна функція ВП має вигляд:

$$\psi\left(\frac{x}{s}\right)=\frac{\text{sign}(x)}{\sqrt{2\pi}}\cdot\exp\left(-\frac{x^2}{2s^2}\right),$$

де s – масштабний коефіцієнт.

Така базисна функція ВП на основі «розщепленої» функції Гауса (ВПРГ) задовольняє ряд умов, що висуваються до такого класу функцій:

- умові локалізації;
- умові допустимості: Фур'є-образ материнського вейвлета $C_\psi=\int\limits_{-\infty}^{+\infty}\frac{|\tilde{\psi}(\omega)|^2}{|\omega|}d\omega<\infty$;
- умові осциляції (знакозмінності) $\int\limits_{-\infty}^{+\infty}\psi(x)dx=0$;
- умові обмеженості: норма базисної функції $\|\psi(x)\|\overset{\Delta}{=}\sqrt{\int\limits_{-\infty}^{+\infty}|\psi(x)|^2dx}<\infty$.

Особливостями такої функції є обмежена ефективна тривалість як у частотній, так і в просторовій областях та можливість її аналітичного аналізу. Така функція має кращу локалізацію, порівняно з ВП, на основі перетворення Гілберта та кращі властивості щодо підкреслення перепадів інтенсивності, порівняно з ВП, на основі похідних функцій Гауса [5, 6].

Перетворення Фур'є базисної функції:

$$\tilde{\psi}(\omega s)=-j\frac{2}{\sqrt{\pi}}s\cdot\exp(-\frac{(\omega s)^2}{2})\cdot\int\limits_0^{\frac{\omega s}{\sqrt{2}}}\exp(z^2)dz\,.$$

Модуль амплітудно-частотної характеристики (АЧХ) цієї функції:

$$\left|\tilde{\psi}(\omega s)\right|=\frac{2}{\sqrt{\pi}}s\cdot\exp(-\frac{(\omega s)^2}{2})\cdot\int\limits_0^{\frac{\omega s}{\sqrt{2}}}\exp(z^2)dz$$

має екстремум у точці:

$$\omega_{extr} \approx \frac{1{,}31}{s}.$$

Таким чином, частота, що відповідає максимуму АЧХ, зменшується обернено пропорційно масштабу ВП, що вказує на частотно-селективні властивості вейвлета.

Неперервне ВПРГ моделі рядка напису (2) має вигляд:

$$WU = \int_{-\infty}^{\infty}\left(U(t-x)\cdot\frac{\operatorname{sign}(t)}{\sqrt{2\pi}}\cdot\exp\left(-\frac{t^2}{2s^2}\right)\right)dt =$$

$$= \begin{cases} s\cdot\sum_{i=1}^{i=N}\left[-\Phi_0\left(\frac{x-A}{s}\right)+\Phi_0\left(\frac{x-A-d-\delta_i}{s}\right)\right], & x < A, \\ s\cdot\sum_{i=1}^{i=N}\left[\Phi_0\left(\frac{x-A}{s}\right)+\Phi_0\left(\frac{x-A-d-\delta_i}{s}\right)\right], & A \le x \le A+d+\delta_i, \\ s\cdot\sum_{i=1}^{i=N}\left[\Phi_0\left(\frac{x-A}{s}\right)-\Phi_0\left(\frac{x-A-d-\delta_i}{s}\right)\right], & x > A+d+\delta_i, \end{cases}$$

де $A = x_0 + (i-1)T + \sum_{k=2}^{i}\varepsilon_{Tk}$;

$\Phi_0(x) = \frac{1}{\sqrt{2\pi}}\cdot\int_0^x \exp\left(-\frac{z^2}{2}\right)dz$ – функція Лапласа.

Дискретизація базисних функцій ВПРГ. При обробці зображень, що зберігаються в цифровому форматі, кожний рядок є дискретизованим варіантом напису $U(k\Delta x)$ ($k = \overline{1,K}$, K – кількість пікселів у рядку зображення), де крок дискретизації $\Delta x = 1$. У цьому випадку і вейвлет $f(m\Delta x_w)$ ($m = \overline{1,M}$, M – кількість дискретних відліків вейвлета), що використовується для аналізу, є результатом дискретизації неперервного вейвлета $f(x)$ з кроком дискретизації Δx_w. Значення *k*-го відліку функції $f(x)$ з кроком $\Delta x_s = \frac{\Delta x_w}{s}$ і масштабованої функції з кроком Δx_w збігаються, тобто $f(k\Delta x_s) = f\left(k\frac{\Delta x_w}{s}\right)$.

Таким чином, зміна масштабу *s* неперервного вейвлета з наступною дискретизацією з кроком Δx_w еквівалентна дискретизації початкового вейвлета з новим кроком $\Delta x_s = \frac{\Delta x_w}{s}$. Одержаний в результаті дискретний вейвлет є масштабованим початковим вейвлетом. При такому підході до визначення масштабу дискретний вейвлет є апроксимацією неперервного вейвлета і збігається з ним при $s \to \infty$ ($\Delta x_s \to 0$), а аналогом масштабу s виступає кількість відліків імпульсної характеристики, тобто довжина дискретного вейвлет-фільтра «розщепленої» функції Гауса (ДВФРГ).

На рис. 2, *в–е* наведено результати ВПРГ рядка зображення з СІ та відповідної моделі СІ (2). При збільшенні масштабу ДВФРГ відбувається зменшення впливу окремих символів та підкреслення границь напису.

Проаналізовано функцію енергії рядка зображення в просторі ВПРГ, яка визначається виразом:

$$En(s) = \int_{-\infty}^{\infty}(WU)^2\,dx.$$

Приклад нормованої залежності функції енергії рядка СІ у просторі ВП наведений на рис. 3.

Для фрагмента напису в рядку зображення функція енергії сигналу в просторі ВПРГ має дві характерні (екстремальні) точки. Локальний максимум відповідає «резонансному» масштабу ВПРГ і розміру символів напису, а локальний мінімум – масштабу, при якому напис сприймається як однорідний об'єкт. При подальшому збільшенні масштабу вигляд функції енергії для фрагмента напису і об'єкта збігається і має монотонно зростаючий характер. Тобто її поведінка стає подібною функції енергії сигналу однорідного об’єкта у просторі ВПРГ.

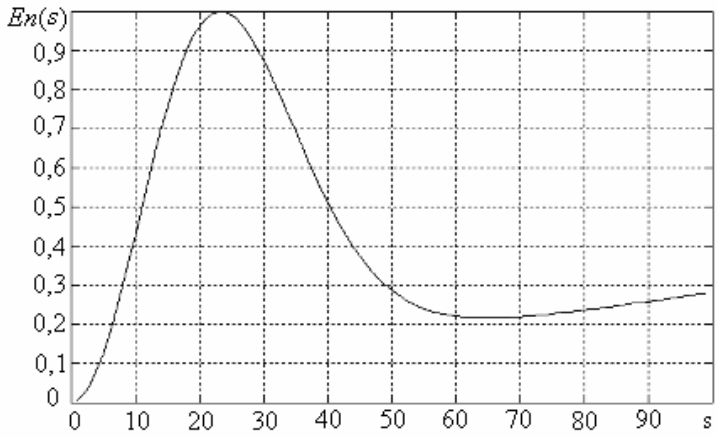


Рис. 3. Залежність енергії рядка зображення з СІ в просторі ВПРГ від масштабу ВП

Аналіз значень екстремумів функції енергії сигналу в просторі ВПРГ дозволяє отримати інформацію про значення масштабу ВП, які необхідно обрати для аналізу структури зображення, положення СІ і розмірів окремих символів, таким чином забезпечуючи адаптивну локалізацію СІ.

Розробка методики адаптивної локалізації. Залежність енергії сигналу від масштабу ВП показує, що у разі, якщо довжина ДВФРГ пропорційна розміру шрифту, то енергія концентрується в області границь окремих елементів символів, що відповідає локальному максимуму функції енергії. Якщо довжина ДВФРГ пропорційна розміру напису, то енергія концентрується в області границі напису, тобто за рахунок зменшення роздільної здатності напис сприймається як однорідний об'єкт. Це відповідає мінімальному значенню залежності. Тобто аналіз максимального та мінімального значень дозволяє одержати інформацію про розташування границь напису та розмір використаного шрифту, що в подальшому дозволяє виконати сегментацію напису на окремі символи.

На базі встановленої залежності енергії рядка з СІ у просторі ВПРГ від масштабу ВПРГ розроблено методику адаптивної локалізації СІ (рис. 4), яка дозволяє визначити інформацію про границі області СІ на зображенні (при обробці ДВФРГ з великим масштабом) та геометричні характеристики символів (при обробці ДВФРГ з «резонансним» масштабом), що може використовуватись для оцінки параметрів моделі формування образів об'єктів початкового зображення.

Рис. 4. Схема методики адаптивної локалізації символьних написів на зображенні з використанням ВП

Запропонована методика адаптивної локалізації СІ містить такі кроки:

– зображення по рядках піддається згортці з імпульсною характеристикою ДВФРГ на різних масштабах – від найбільшого до найменшого;
– відбувається вибірковий аналіз функції енергії рядка зображення та визначаються потрібні масштаби;
– визначаються границі однорідних областей;
– визначають характеристики геометричних розмірів символів;
– обробка здійснюється послідовно по всіх рядках.

На рис. 5 наведено результати комп'ютерної обробки зображення, що містить СІ (рис. 1), які демонструють роботу даної методики адаптивної локалізації СІ та підтверджують можливість виділення областей написів з різними шрифтами.

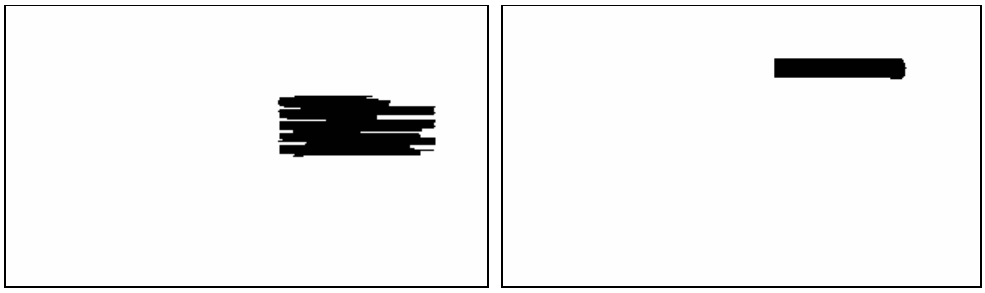


Рис. 5. Приклади виділення областей символьних написів з різними розмірами шрифту на зображенні (рис. 1)

Висновки. У статті розроблено метод адаптивної локалізації символьної інформації на зображенні. Для визначення областей з СІ запропоновано враховувати як інформацію про її границі, так і геометричні характеристики окремих символів. Така інформація може бути визначенна шляхом аналізу функції енергії рядка символьного напису в просторі ВПРГ. Запропонований метод локалізації СІ дозволив отримати параметри моделі формування образів об'єктів на зображеннях з СІ. Використання методики адаптивної локалізації дозволило підвищити вірогідність розпізнавання ієрархічно складно організованих об'єктів, що містять СІ у різних практичних додатках.

ЛІТЕРАТУРА:

1. *Гонсалес Р., Вудс Р.* Цифровая обработка изображений. – М.: Техносфера, 2005. – 1072 с.
2. MPEG-7 – http://drogo.cselt.it/mpeg.
3. *Антощук С.Г., Бабилунга О.Ю., Николенко А.А.* Иерархическая модель при формировании образов объектов // Вестник Национального технического университета "ХПІ": Сб. научных трудов. – Тематический выпуск "Информатика и моделирование". – 2007. – № 39. – С. 3–12.
4. *Pratt W.K.* Digital image processing. – 3nd ed. – USA: John Wiley & Sons, 2001. – 750 p.
5. *Добеши И.* Десять лекций по вейвлетам. – М.: Ижевск: РХД, 2001. – 406 с.
6. *Клих Ю.А., Антощук С.Г., Николенко А.А.* Адаптивные базисные функции вейвлетного преобразования // Труды Одес. политехн. ун-та. – Одесса, 2004. – Вып. 2 (22). – С. 121–125.

АНТОЩУК Світлана Григорівна – доктор технічних наук, професор, завідувач кафедри інформаційних систем Одеського національного політехнічного університету.

Наукові інтереси:
- системи штучного інтелекту;
- інформаційні технології.

БАБІЛУНГА Оксана Юріївна – кандидат технічних наук, доцент кафедри інформаційних систем Одеського національного політехнічного університету.

Наукові інтереси:
- розпізнавання образів;
- цифрова обробка зображень.

НІКОЛЕНКО Анатолій Олександрович – кандидат технічних наук, доцент кафедри інформаційних систем Одеського національного політехнічного університету.

Наукові інтереси:
- теорія вейвлет-перетворень;
- цифрова обробка сигналів.

ТКАЧЕНКО Олена Вікторівна – асистент кафедри інформаційних систем Одеського національного політехнічного університету.

Наукові інтереси:
- нейромережеві технології;
- системи штучного інтелекту.

Подано 15.11.2008

Антощук С.Г., Бабілунга О.Ю., Ніколенко А.О., Ткаченко О.В. Адаптивна локалізація символьних написів на зображеннях методом вейвлет-аналізу

Антощук С.Г., Бабилунга О.Ю., Николенко А.О., Ткаченко О.В. Адаптивная локализация символьных надписей на изображениях методом вейвлет-анализа

Antoschuk S.G., Babilunga O.Y., Nikolenko A.O., Tkachenko O.V. Adaptive localization of character inscriptions on images offered approach method of veyvlet-analysis

УДК 658.5.011.56:004.932.001.57

Адаптивная локализация символьных надписей на изображениях методом вейвлет-анализа / С.Г. Антощук, О.Ю. Бабилунга, А.О. Николенко, О.В. Ткаченко

Предложен подход к адаптивной локализации символьной информации на изображениях путем анализа функции энергии сигнала в просторные вейвлет-превращения, что позволило разработать модель формирования образов объектов на изображении и соответствующие информационные технологии.

УДК 658.5.011.56:004.932.001.57

Adaptive localization of character inscriptions on images offered approach method of veyvlet-analysis / S.G. Antoschuk, O.Y. Babilunga, A.O. Nikolenko, O.V. Tkachenko

To adaptive localization of character information on images by the analysis of function of energy of signal in spacious veyvlet-transformations, that allowed to develop the model of forming of appearances of objects dark-and-light and proper information technologies.