

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

УДК 621.372.542

В.П. МАЙДАНЮК, В. І. МЕСЮРА, А. О. СТАХОВ

Вінницький національний технічний університет, Вінниця

КОДУВАННЯ РУХОМИХ ЗОБРАЖЕНЬ НА ОСНОВІ ФОРМУВАННЯ ПРОСТОРОВИХ ЧАСТОТНИХ КОМПОНЕНТ

Анотація: Розглянуто кодування зображень на основі адаптивного до контурів двовимірного аналізу і синтезу, зокрема, застосування його до кодування відео зображень. *Запропоновано відновлення відліків низькочастотних компонент методами інтерполяції і контурного передбачення, що призводить до збільшення візуальної якості зображення при тій же "детальності", а також схему детектора руху, що забезпечує можливість передачі високочастотної складової відео зображення методом чергування.*

Ключові слова: Відео, детектор руху, зображення, лінійна інтерполяція, кодування, контур, передбачення.

Вступ

З появою стандарту MPEG [1-2] для кодування відео в тіні залишились інші методи кодування. До MPEG виконувались інтенсивні дослідження цифрових методів передачі відео зображень, які однак не знайшли широкого застосування, оскільки були витіснені цим стандартом, незважаючи на його недоліки. Зокрема, стандарт MPEG вимагає доволі потужних обчислювальних засобів невідповідних якості відновленого зображення.

З цієї точки зору заслуговує на увагу метод кодування зображень, відомий під назвою «адаптивний до контурів двовимірний аналіз і синтез зображень». Цей метод раніше інтенсивно досліджувався для застосування в системах цифрового телебачення [3-5]. Його особливістю є формування декількох низькочастотних та високочастотних сигналів, які несуть інформацію про деталі зображення різних розмірів, що відкриває широкі можливості маніпулювання ними при передачі відео. Зокрема, з'являються можливості застосування міжкадрової інтерполяції та детектування руху, передачі відліків високочастотних компонент з чергуванням і т.п., що може значно знизити цифрові потоки при передачі відео.

Актуальність

Цифрові потоки і якість відновленого зображення при адаптивному до контурів аналізі і синтезі в значній мірі залежать від точності формування низькочастотних компонент зображень та точності детектування руху.

Проте існуючі підходи, які ґрунтуються на лінійних методах формування відліків низькочастотних компонент, не враховують особливостей сприйняття зображення людиною і обмежують можливості методу. Тому розгляд цих питань є актуальним.

Мета

Метою статті є вдосконалення методів формування низькочастотних компонент зображення та детектування руху при адаптивному до контурів двовимірному аналізі і синтезі.

Задачі

Відповідно до мети досліджень формулюються такі задачі:

- 1) Аналіз методу формування низькочастотних компонент зображення та його вдосконалення.
- 2) Розробка схеми детектування руху.

Модель адаптивного до контурів двовимірного аналізу і синтезу зображень

У техніці телебачення звичайний процес перетворення сигналу полягає в почерговій, циклічно повторюваній передачі відліків яскравості окремих елементів зображення. При цьому аналіз зображення здійснюється однією апертурою, що має позитивну функцію прозорості, симетричну відносно осі апертури. Така апертура утворює двовимірний низькочастотний фільтр, так званої, просторової частоти.

Збільшення розмірів апертури приводить до зникнення деталей, розміри яких менше ефективного діаметру апертури, і відповідно, до збільшення нерізкості, що еквівалентно скороченню смуги пропускання двовимірного фільтра. Ефективний діаметр апертури визначає кількість елементів зображення і розмір одного елемента.

У даному способі аналізу та синтезу відео зображення враховується розмір деталей у двох вимірах, що дозволяє шляхом більш повного врахування особливостей зорового сприйняття підвищити економічність відео систем, передавати відео повідомлення по каналах з зменшеною пропускну здатністю без помітного погіршення якості, спростити процеси запису і відтворення відео сигналу. Передане повідомлення, тобто відео зображення $V(x,y)$, піддається одночасному впливу кількох, які відрізняються один від одного аналізуючих апертур, які можуть мати як ділянки з "позитивною прозорістю", так і ділянки з "негативною прозорістю". Сукупність таких апертур дозволяє виділити на

різних виходах аналізатора сигнали, які несуть інформацію про деталі різної величини, а разом з тим, звичайно і про відповідні градації контурних обрисів переходів при передачі контрастних кордонів.

З точки зору двовимірного перетворення Фур'є сукупність зазначених апертур становить групу двовимірних фільтрів так званої просторової частоти. Розмір елемента зображення, отриманого після синтезу, визначається апертурою, яка дорівнює сумі функцій розподілу окремих апертур. Найбільша апертура має позитивну прозорість і являє собою двовимірний фільтр низької частоти. Сигнал, який створюється цією апертурою, дозволяє побудувати зображення, відповідне дефокусованому вихідному зображенню. При цьому число ефективно помітних елементів зображення зменшується пропорційно квадрату ефективного діаметру апертури. Апертури, що володіють ділянками позитивної та негативної прозорості можна розглядати як результат віднімання двох апертур з різними ефективними діаметрами, з суміщеними осями і з рівними "обсягами функції розподілу". З точки зору двовимірного перетворення Фур'є такі апертури представляє собою деякий двовимірний смуговий фільтр просторової частоти, межами смуги пропускання якого є характеристики двовимірних фільтрів низької частоти, що утворили даний смуговий фільтр.

Для досягнення необхідної точності аналізу та синтезу необхідно забезпечити правильне примикання смуг пропускання двовимірних фільтрів. Іншими словами, верхня межа смуги пропускання одного фільтра повинна з достатньою точністю відповідати нижній межі смуги пропускання сусіднього фільтра. При цьому функції розподілу апертур, що здійснюють аналіз вихідного зображення, у сумі утворюють апертуру, функції розподілу і ефективний діаметр якої були задані при виборі стандарту чіткості відео системи. Чіткість синтезованого зображення при виконанні умови правильного примикання відповідає чіткості, яку забезпечує звичайний спосіб аналізу та синтезу зображення однієї апертурою з таким же ефективним діаметром.

З огляду на те, що аналіз зображення ведеться одночасно декількома апертурами, що виконують функції двовимірних фільтрів, вихідне зображення створює одночасно кілька сигналів, що несуть інформацію про деталі різної величини, причому при поділі інформації розміри деталей враховані у двох вимірах. Це дозволяє більш правильно вибрати необхідну точність відтворення шляхом більш повного використання особливостей зорового сприйняття. При обробці зазначених сигналів передбачається не тільки відомий прийом квантування окремих сигналів на різне число рівнів, що використовує особливості контрастної чутливості, але також використання інерційності зору шляхом застосування квантування з змінними за часом порогоми квантування. Цей прийом дозволяє при заданому числі рівнів квантування поліпшити якість, зменшивши вплив шуму квантування або при заданій якості зменшити кількість рівнів квантування і скоротити обсяг відео сигналу.

Процес розкладання на двовимірні компоненти можна представити таким чином. Нехай $\Phi_1, \Phi_2, \dots, \Phi_n$ низькочастотні компоненти, отримані низькочастотною фільтрацією вихідного зображення. Причому, смуга частот $W_1 > W_2 > \dots > W_n$. Тоді процес розкладання на компоненти можна представити наступним виразом:

$$\tilde{\Phi} = \Phi_n + (\Phi_{n-1} - \Phi_n) + (\Phi_{n-2} - \Phi_{n-1}) + \dots + (\Phi_1 - \Phi_2) + (\Phi_0 - \Phi_1). \quad (1)$$

Розкривши дужки, одержимо

$$\tilde{\Phi} = \Phi_0$$

Тобто не враховуючи спотворень, які вносяться апаратурою, таке розкладання дозволяє однозначно відновити вихідний сигнал. Процес кодування полягає в дискретизації кожної компоненти відповідно до теореми Котельникова і квантуванні на число рівнів відповідне особливостям сприйняття даної смуги частот зоровим аналізатором людини. В лінію зв'язку передаються відліки компоненти Φ_n і відліки різницевих компонент [4].

Лінійна інтерполяція і контурне передбачення при двовимірному аналізі і синтезі

Особливістю адаптивного до контурів двовимірного аналізу і синтезу зображень є формування декількох сигналів, кожен з яких відповідає зображенню зі збільшеною нерізкістю з подальшим квантуванням на різне число рівнів попарних різниць цих компонент. Від точності формування відліків низькочастотних компонент залежить якість відновленого зображення і коефіцієнт стиснення.

Найбільш економічним з точки зору технічних витрат є формування згладжених компонент за рахунок зниження частоти дискретизації у кожній наступній компоненті в два рази. Це аналогічно обчисленню середнього значення в межах апертур, що примикають одна до одної з розмірами кратними 2^n . У результаті такої операції формується поле рідких відліків. Природно припустити, що у випадку

лінійного згладжування вихідного зображення пропущені відліки необхідно також відновлювати лінійними інтерполяційними методами. Однак, застосування лінійних інтерполяційних методів не дозволяє точно відновлювати контурні переходи на зображенні.

Одним із способів підвищення ефективності кодування при двовимірному аналізі і синтезі є застосування контурного передбачення при відновленні відліків низькочастотних компонент. Дійсно основна інформація про зображення може бути отримана з аналізу контурів цього зображення. Тому було б логічно використовувати цю властивість при кодуванні (ущільненні) зображень.

Суть цього методу полягає в наступному. На ділянках зображення з плавною зміною яскравості проміжні відліки відновлюються лінійною інтерполяцією, а на ділянках зображення з різкою зміною яскравості - передбаченням по контуру. До контуру повторюють попереднє значення яскравості, а після контуру наступне. Це дозволяє більш точно відновити контурні переходи низькочастотних компонент, що сприяє зменшенню числа розрядів різницевого сигналу.

Це добре ілюструє рис. 1, на прикладі одновимірного сигналу, де a - це початковий сигнал, b - відновлений інтерполяцією по опорних відліках, а c - відновлений з застосуванням як інтерполяції на плавних ділянках, так і контурного передбачення на ділянках зі стрибкоподібним зміною значень яскравості.

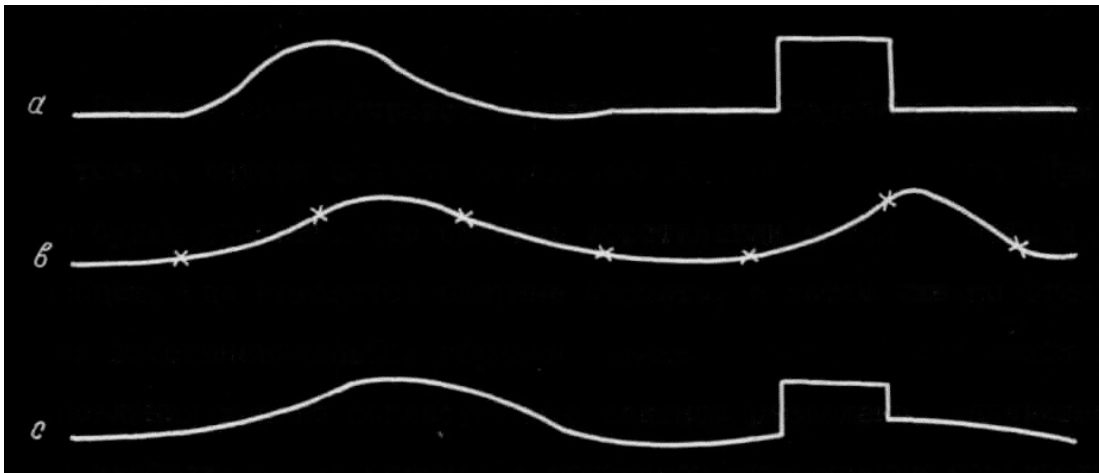


Рисунок 1 – Відновлення сигналу інтерполяцією і передбаченням по контуру

Нескладно побачити, що при відновленні сигналу тільки лінійними інтерполяційними методами (рис. 1) прямокутний імпульс перетворюється на трикутний, а при відновленні методами контурного передбачення форма сигналу зберігається. Спектральна щільність прямокутного імпульсу визначається так [12]:

$$S_1(\omega) = At_u \left[\frac{\sin \omega t_u / 2}{\omega t_u / 2} \right] \quad (2)$$

А спектральна щільність трикутного імпульсу така [12]:

$$S_1(\omega) = \frac{At_u}{2} \left[\frac{\sin \omega t_u / 4}{\omega t_u / 4} \right]^2 \quad (3)$$

Звідки видно, що спектральна щільність прямокутного імпульсу зменшується пропорційно $1/\omega$, а трикутного - пропорційно $1/\omega^2$. Тобто при лінійній інтерполяції придушуються високочастотні складові сигналу, що стосовно до зображення призводить до "розмивання" контурів. А це в свою чергу, до зменшення візуальної чіткості зображення. Тому на ділянках зображення з плавною зміною яскравості пропущені відліки відновлюються лінійною інтерполяцією, а на ділянках зображення з різкою зміною яскравості – передбаченням по контуру.

Позначимо опорні відліки (середні значення яскравості в групах $N \times M$), як $X(0,0)$, $X(0,1)$, $X(1,0)$, $X(1,1)$ (рис. 2). Тоді на ділянках з плавною зміною яскравості для відновлення пропущених відліків необхідно підсумувати крайні елементи апертури кожен зі своєю вагою в залежності від координати точки в просторі зображення.

Для обчислення шуканого відліку при лінійній одновимірній інтерполяції використовуються тільки два послідовних відліки вхідної послідовності $X(n)$, причому значення інтерпольовані між двома відліками $X(0)$, $X(1)$, лежать на прямій лінії, що сполучає два вхідних відліки. Таким чином, співвідношення, що зв'язує вихідний сигнал інтерполятора і вхідну послідовність $X(n)$ для одновимірного випадку має такий вигляд (вздовж осі X , "0" рядок):

$$Y(0, n_2) = X(0,0) + \frac{X(0,1) - X(0,0)}{M} n_2 = \frac{(M - n_2)X(0,0) + n_2 X(0,1)}{M}, \quad (4)$$

$$0 \leq n_2 \leq M$$

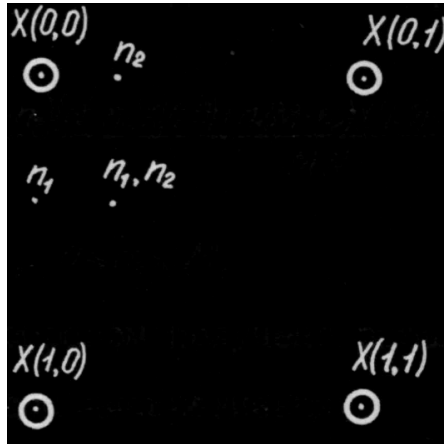


Рисунок 2 – Розташування опорних відліків в просторі

Так само може бути визначена будь-яка точка і в напрямку осі y :

$$Y(n_1, 0) = X(0,0) + \frac{X(1,0) - X(0,0)}{N} n_1 = \frac{(N - n_1)X(0,0) + n_1 X(1,0)}{N}, \quad (5)$$

$$0 \leq n_1 \leq N$$

Фільтр-інтерполятор залишає незмінним спектральні складові сигналу з частотами менше π / M в напрямку X і π / N в напрямку Y і подавляє спектральні складові з великими частотами.

Тому у разі наявності контуру між опорними відліками до контуру потрібно повторити попередній опорний відлік, а після контуру - наступний. Наприклад, для 0-рядки попереднього прикладу це можна записати наступним чином:

$$Y(0, n_2) = \begin{cases} X(0,0) & \text{— до контура} \\ X(0,1) & \text{— після контура} \end{cases} \quad (6)$$

У загальному випадку відновлену низькочастотну компоненту можна представити як результат обробки дискретизованого низькочастотного сигналу оператором лінійної інтерполяції і оператором передбачення:

$$Y = AX + BX \quad (7)$$

Причому A визначений на неперервних ділянках зображення і є оператором лінійної інтерполяції, а B - оператор контурного передбачення і визначений на ділянках зображення із стрибкоподібним зміною яскравості зображення.

Таким чином, застосування інтерполяції і контурного передбачення дозволяє, з одного боку усунути "помилкові" контури на ділянках зображення з плавною зміною яскравості, а з іншого боку - відновлює крутизну перехідної характеристики в місцях наявності контуру, що призводить до збільшення візуальної якості зображення при тій же "детальності".

Детектування руху та передача високочастотної складової зображення

Присутність в послідовності відео кадрів рухомих об'єктів, викликає міжкадрові зміни яскравості, якими можна скористатися для виявлення руху і оцінки параметрів руху. Безпосереднє віднімання двох послідовних кадрів початкового зображення не може служити оцінкою наявності чи відсутності руху, так як при цьому виділяються шуми, які можуть давати помилкові ознаки руху, до того ж, необхідна велика пам'ять на кадр.

Але ситуація істотно змінюється, якщо беруться міжкадрові різниці відліків низькочастотної компоненти, наприклад компоненти H_3 , яка визначена середніми значеннями в межах примикаючих один до одного фрагментів зображення з розмірами 8×8 , або компоненти H_2 (у разі відсутності компоненти H_3), яка визначається середніми значеннями по фрагментах 4×4 . По-перше, за рахунок усереднення рівень шуму тут значно знижений, а по-друге, для організації кадрової пам'яті потрібно набагато менше витрат обладнання (об'єм пам'яті в 64 (16) рази менше). Таким чином наявність в переданому сигналі низькочастотних компонент дозволяє без значних додаткових апаратурних витрат реалізувати детектор руху, що забезпечує можливість передачі високочастотної складової відео зображення методом чергування. Тобто в кожному кадрі передається лише частина відліків високочастотної компоненти. Наприклад, 1-й кадр – непарні рядки, непарні відліки; 2-й кадр – парні рядки, парні відліки; 3-й кадр – непарні рядки, парні відліки; 4-й кадр – парні рядки, непарні відліки.

На приймальній стороні на нерухомих ділянках відтворюється повна роздільна здатність за рахунок накопичення в кадровому запам'ятовуючому пристрої відліків попередніх кадрів високочастотної складової, а на рухомих ділянках високочастотна складова виключається з процесу синтезу зображення, але візуальна якість відео зображення не зменшується, оскільки роздільна здатність зорового аналізатора людини також зменшується при спостереженні рухомих об'єктів.

Структурна схема детектора руху наведена на рис. 3, де L_k – затримка на кадр зображення, P – пороговий пристрій, X_k - значення відліків низькочастотної компоненти поточного кадру, W_k, W_{k-1} - проміжні вихідні сигнали рекурсивного фільтра, Y_k - вихідний різницевий сигнал, M – сигнал руху.

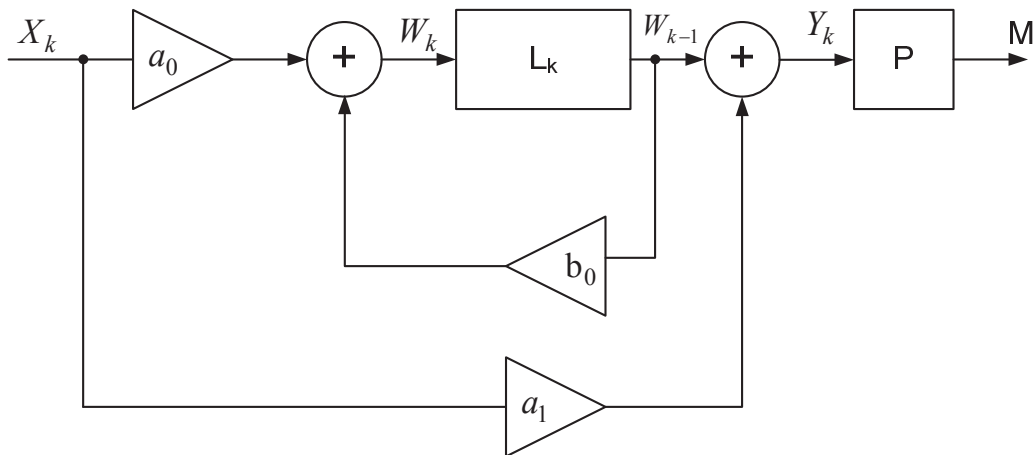


Рисунок 3 – Структурна схема детектора руху

Визначимо перехідну характеристику цього пристрою. Різницеві рівняння для проміжного виходу W_k та різницевого сигналу Y_k визначаються наступними виразами:

$$\begin{aligned} W_k &= X_k a_0 + W_{k-1} b_0 \\ Y_k &= X_k a_1 + W_{k-1} \end{aligned} \quad (8)$$

Виконавши Z – перетворення рівнянь 8, знаходимо:

$$\begin{aligned} W(Z) &= X(Z) a_0 + W(Z) Z^{-1} b_0 \\ Y(Z) &= X(Z) a_1 + W(Z) Z^{-1}. \end{aligned} \quad (9)$$

Звідки

$$Y(Z) = a_1 X(Z) + \frac{X(Z)a_0 Z^{-1}}{1 - b_0 Z^{-1}}. \quad (10)$$

Для визначення реакції пристрою на одиничну функцію у вираз (10) замість $X(Z)$ підставимо значення

$$X(Z) = \frac{1}{1 - Z^{-1}},$$

яке є Z -перетворенням одиничної функції

$$X(t) = \begin{cases} 1, & t \geq 0, \\ 0, & t < 0. \end{cases}$$

Тоді

$$Y(Z) = \frac{a_1}{1 - Z^{-1}} + \frac{a_0}{1 - b_0} \left(\frac{1}{1 - Z^{-1}} - \frac{1}{1 - b_0 Z^{-1}} \right). \quad (11)$$

Взявши зворотнє Z -перетворення від виразу (11), отримаємо:

$$Y(k) = \frac{a_1(1 - b_0) + a_0(1 - b_0^k)}{1 - b_0}. \quad (12)$$

Даний вираз визначає перехідну характеристику детектора руху, k – номер кадру, $k=0$ – поточний кадр. Нехай $a_1 = -1$, $a_0 = b_0 = 1/2$, тоді неважко визначити, що така перехідна характеристика забезпечить ознаку руху на протязі 4-х кадрів, що відповідає глибині чергування високочастотної компоненти. Це необхідно для виключення високочастотної складової з процесу синтезу зображення на приймальній стороні після виникнення руху до повної її зміни.

Висновки

1. Стандарт MPEG вимагає доволі потужних обчислювальних засобів невідповідних якості відновленого зображення, тому потрібні інші методи ущільнення зображень з втратами, які характеризувались би меншими обчислювальними затратами.

2. Серед таких методів заслуговує на увагу метод кодування зображень, відомий під назвою «адаптивний до контурів двовимірний аналіз і синтез зображень». Його особливістю є формування декількох низькочастотних та високочастотних сигналів, які несуть інформацію про деталі зображення різних розмірів, що відкриває широкі можливості маніпулювання ними при передачі відео. Зокрема, з'являються можливості застосування міжкадрової інтерполяції та детектування руху, передачі відліків високочастотних компонент з чергуванням і т.п., що може значно знизити цифрові потоки при передачі відео.

3. Одним із способів підвищення ефективності кодування при двовимірному аналізі і синтезі є застосування інтерполяції і контурного передбачення при відновленні відліків низькочастотних компонент, що відновлює крутизну перехідної характеристики в місцях наявності контуру і призводить до збільшення візуальної якості зображення при тій же "детальності".

4. Наявність в переданому сигналі низькочастотних компонент дозволяє без значних додаткових апаратних витрат реалізувати детектор руху, що забезпечує можливість передачі високочастотної складової відео зображення методом чергування.

Список використаної літератури

- 1.Международный стандарт MPEG (ISO/IEC 14496).
- 2.Гонсалес Р., Вудс Р. Цифровая обработка изображений.- М.: Техносфера, 2005. – 1072 с.
- 3.Брауде-Золотарем Ю.М. Исследование возможностей сокращения объема телевизионного сигнала за счет использования свойств зрения: Автореф. дис. канд.тех. наук. – М., 1960. – 15 с.
- 4.Майданюк В.П. Разработка алгоритмов и аппаратных средств систем сжатия телевизионных изображений: Автореф. канд.тех. наук. – Винница, 1993. – 22 с.

5. Кожем'яко В. П., Майданюк В. П., Теренчук А. Т., Хіллес Шаді Мазін. Адаптивний до контурів двовимірний аналіз і синтез // Міжнародний науково-технічний журнал «Оптико-електронні інформаційно-енергетичні технології, 2002. - № 4. - С. 120-127.

Відомості про авторів

Майданюк Володимир Павлович – к. т. н., доцент кафедри програмного забезпечення; Хмельницьке шосе, 95, м. Вінниця, 21021; тел. 0965636971; e-mail: maydan2000@mail.ru, Вінницький національний технічний університет;

Месюра Володимир Іванович – к. т. н., професор кафедри інтелектуальних систем; Хмельницьке шосе, 95, м. Вінниця, 21021; тел. +0972647731; e-mail: vimes2009@yandex.ru, Вінницький національний технічний університет.

Стахов Артем Олександрович – магістр кафедри програмного забезпечення; Хмельницьке шосе, 95, м. Вінниця, 21021. Вінницький національний технічний університет.