

УДК 004.021

Юрко О.В.

Криворізький національний університет

Особливості роботи із зображеннями великих розмірів (ЗВР) та ієрархічна обробка.

Розглядається проблематика роботи з зображеннями великих розмірів, та запропоновано комбінований рекурсивний алгоритм кодування зображень на основі декомпозиції квадродеревом деревом та фільтрації з метою подолання цих проблем.

Ключові слова – великі зображення, кодування зображень, рекурсивний алгоритм, формат зображень.

На теперішній час прогрес технології та науки в виробництві світло чутливих матриць призвів до значного здешевлення і як слідство поширення пристроїв які можуть створювати цифрові зображення великої роздільної здатності. Так наприклад провідні виробники фототехніки такі як Sony, Canon, Nikon починаючи з 2007-2008 років серійно виробляють для споживчого ринку фотоапарати з CMOS матрицями роздільною здатністю від 21,1 до 24,8 мегапікселей. А фірма Leica серійно випускає фотоапарат Leica S2 з сенсором 37,5 мегапіксела. Виробники сучасного медичного обладнання використовують в рентгенівських апаратах датчики з роздільною здатністю до 10-20 ліній на 1мм [1], що на стандартний розмір рентгенівської фотографії 40x40 см складає 16-64 мегапікселей. Комп'ютерні томографи дозволяють формувати зображення з роздільною здатністю до 0,35 мм. [2]

Людина більш ніж 90% інформації про навколишнє середовище отримує через органи зору. А в професійних сферах можна сказати що 99%. С розвитком науки і техніки особливе значення стало мати застосування великих растрових зображень в інформаційних системах. Растрових, тому що саме отриманий растр є первинною інформацією яка слугує для подальшої обробки, та переведення в векторне зображення, яке по суті є певною моделлю об'єкта, а процес та алгоритм цього переведення залежить від предметної області. На сьогодні нема принципових технічних обмежень для розмірів растрових зображень. Однак робота із растровими зображеннями великих розмірів суттєво відрізняється від обробки звичних зображень малих форматів. Відмінності лежать в площині отримання зображень, зберігання, передачі й сприйняття оператором. Ігнорування цих відмінностей створює цілу низку проблем пов'язаних із ефективністю обробки й застосування великих зображень. Найбільш серйозною проблемою є передача таких зображень через радіотехнічні засоби зв'язку. Наприклад із безпілотних літальних апаратів.

В даній роботі аналізуються методи й засоби отримання ЗВР, зберігання, передачі та сприйняття оператором, пропонується методика роботи із ЗВР (зберігання, передача, сприйняття) з ієрархічним багаточастотним розкладом і синтезом.

Людське око є оптичною системою яка створює та передає зображення у мозок для подальшої обробки. Саме око з інженерної точки зору послугувало спочатку прототипом для розробки оптичних пристроїв, а потім і алгоритмів обробки зображень. Людське око містить близько 120 мільйонів паличок та 6 мільйонів колбочок, тобто можна сказати що в загалі апарат зору має роздільну здатність 250 мільйонів пікселів, якщо використовувати технічну термінологію. Але треба відмітити що палички і колбочки в оці розташовуються не рівномірно. Навколо так званої ямки, яка знаходиться на оптичній вісі ока розташування найбільш щільне, і там знаходяться майже всі колбочки, які є сенсорами кольору, і роздільна здатність в цій області найкраща, а на периферії ока роздільна здатність набагато гірша, палички об'єднуються в більші області, які по суті виконують функції бокового бачення та «детектора руху». Таким чином можна вважати що людина при роботі з важливою інформацією використовує лише частину з наданої природою кількості сенсорів. Це палички і колбочки розташовані у ямці сітківки ока. Але це не є недоліком, так як завдяки можливості швидко позиціонувати увагу, або оптичну вісь ока на важливих елементах, що дає можливість мати представлення про оточуючу дійсність з бажаною деталізацією. Також при такому підході спрощується подальша обробка отриманого зображення, так як в один і той же момент часу, аналізується тільки не велика частина оточуючої дійсності.

В технічних системах, які працюють з великими зображеннями, проблему підвищення деталізації вирішують за допомогою збільшення роздільної здатності сенсора, використання кількох сенсорів з наступною зшивкою отриманих зображень. У випадку коли необхідно

спостерігати за деякими об'єктами дистанційно, наприклад камери встановленні на безпілотних літаючих апаратах, можуть використовуватися камери не великої роздільної здатності, але ці камери розташовуються на рухомих платформах, та оснащуються оптичним механізмом масштабування, що дозволяє оператору в деталях розглядати об'єкти, які його зацікавили. Але у такого підходу є деякі недоліки. В випадку коли одночасно с системою потрібно працювати кільком операторам, яким потрібні різні об'єкти, необхідно застосовувати кілька камер, та кілька механічних систем позиціювання, по кількості операторів, що не завжди технічно та економічно можливо і доцільно. Також дуже складно, за допомогою такої системи оброблювати інформацію в ретроспективі, так як зберігати можна тільки ті дані на які звернув увагу оператор, і відповідно зпозиціонував камеру. Таким чином великий інтерес представляють системи які забезпечують формування детального зображення за рахунок використання камер, які встановленні без механічних пристроїв наведення та збільшення, а їх відсутність компенсувати використанням сенсорами великої роздільної здатності.

Великі розміри зображень можна отримати використовуючи зшивання менших зображень. Такий підхід широко застосовується для отримання зображень роздільна здатність яких перевищує роздільну здатність окремого сенсору.

Найбільше коло застосування багато сенсорні камери знаходять в аерофотозйомці, зйомці земної поверхні з космічних апаратів та в останній час в астрономічних дослідженнях.

Для аерофотозйомки промисловістю випускається ряд камер роздільною здатністю від 60-91 мегапікселів, так звані середньо форматні камери, до 250 мегапікселів – крупно форматні камери [3]. Така роздільна здатність забезпечується використанням в конструкції камери кількох сенсорів, і кількох об'єктивів. Окремий класом йдуть аерофотосканери, в яких сенсором є не CCD або CMOS матриця, а відповідна лінійка, і за рахунок лінійного переміщення літака отримується зображення маршруту. Такі лінійки мають до 12000 чуттєвих елементів Тобто розмір отриманого зображення залежить від довжини маршруту помноженої на кількість чуттєвих елементів лінійки, та ємністю пристроїв пам'яті [4]



Рис. 1. Зовнішній вигляд цифрових камер для аерофотозйомки

Конструктивно пристрої пам'яті куди зберігаються отримані зображення представляють собою або герметичні блоки жорстких дисків, або твердотільну пам'ять загальною ємністю 1 – 4,5 Тб. Майже все камери підтримують можливість зміни пристроїв пам'яті в польоті. В якості формату зберігання зображень використовуються TIFF (як з стисненням так і без), JPEG, JPEG2000, у окремих камер є можливість зберігати «сирі дані» (RAW).

Інший підхід для вирішення проблеми отримання високо деталізованого зображення застосовується фірмою Gigapan [5]. Нею пропонується програмно апаратний комплекс з тією ж назвою. Принцип формування зображень полягає в тому, що цифрова камера робить серію знімків з перекриттям для наступної зшивки отриманих зображень в одне у програмній частині комплексу. За позиціонування камери, відповідає роботизований пристрій. Цей пристрій наводить камеру на нову точку, та подає команду камері на зйомку.



Рис. 2. Зовнішній вигляд рухомої платформи фірми Gigapan з встановленим цифровим фотоапаратом

Використання високоякісної оптики, з великим збільшенням дозволяє отримати зображення з неперевершеною деталізацією. Саме зображенням створеним за технологією Gigapan належать рекорди по роздільній здатності, та розміру. Прикладом може слугувати панорама міста Шанхай, яка має розмір 272 гігапіксела. Її розмір 887 276 на 306 908 пікселів, а об'єм даних складає 1,09 терабайта. Загальна кількість зроблених окремих знімків, які було зроблені для створення цього зображення складає 12000 [6].

В портфоліо користувачів продукції фірми Gigapan десятки зображень розмір яких перевищує 10 гігапікселів, а об'єм даних десятків гігабайт.

Перевагами такого підходу є доступність, немає потреби використовувати спеціалізовані камери, достатнім є використання звичайних любительських або напівпрофесійних камер. Недоліком такого підходу у створенні зображень є те що зйомка панорам займає дуже багато часу. Тому окремі частини отриманого зображення можуть не совпадати. Це зв'язано з тим, що об'єкти які фотографують можуть переміщуватися. Тому досить часто на таких зображеннях можна спостерігати артефакти на місцях зшивки окремих кадрів.

Але найбільше практичне використання, зображення великих розмірів отримати в аерофотозйомці. Для цих цілей випускаються спеціальні фотоапарати які можуть видавати зображення до 1,8 гігапікселів включно [7] Камера забезпечує таку роздільну здатність завдяки використанню 4 об'єктивів та 368 світлочувливих матриць від мобільних телефонів роздільною здатністю 5 мегапікселів кожен. Також військова агенція США DARPA фінансує розробку 1 гігапіксельної однооб'єктивної камери AWARE-2 [8].

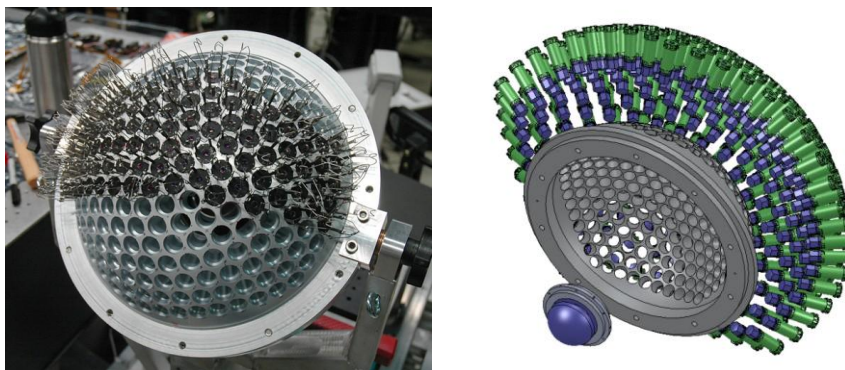


Рис. 3. Камера Aware-2

Як вказують автори, використаний ними підхід при розробці камери AWARE-2 може бути використаний для створення камер з роздільною здатністю 50-100 гігапікселів.

В світі інтенсивно йдуть розробки щодо збільшення роздільної здатності камер в таблиці 1 приведено технічні характеристики діючих експериментальних камер, роздільна здатність яких перевищує 1 гігапіксель.

Табл1. Характеристики гігапіксельних камер

Назва системи	Поле зору, гр.	Роздільна здатність	Об'єм оптичної складової системи	Об'єм системи, м ³
---------------	----------------	---------------------	----------------------------------	-------------------------------

			дм ³	
AWARE-2	120 ⁰ х50 ⁰	1,2х10 ⁹	8х10 ⁻³	3х10 ⁻¹
AWARE-40	120 ⁰ коло	2,8х10 ¹⁰	1х10 ⁻¹	5х10 ⁻¹
Multi-lens array	72 ⁰ коло	4,1х10 ⁸	1х10 ⁻¹	-
Asymmagon	90 ⁰ х45 ⁰	>6.7х10 ⁷	8х10 ⁻³	9х10 ⁻²
ARGUS	60 ⁰ коло	1,4х10 ⁹	3х10 ⁻²	3х10 ⁻¹
LSST	3,5 ⁰ коло	3,3х10 ⁹	3х10 ²	1х10 ⁴
Pan-STARRS PS1	3 ⁰ коло	1,0х10 ⁹	7х10 ⁰	3х10 ³

Таким чином на сьогодні є технічна можливість отримувати зображення роздільна здатність яких обмежена лише носіями інформації для їх зберігання та технологією їх обробки. Також необхідно зазначити, що при збільшенні роздільної здатності зображень об'єм даних зростає в квадратичній залежності. Що теж не є великою проблемою так як сучасні системи зберігання даних побудовані на накопичувачах на жорстких дисках можуть досягати петабайтних розмірів. Проблемою є лінії, що передають дані від сховища даних до кінцевого користувача. Тобто, якщо не прийняти спеціальних заходів необхідно буде передавати терабайти даних, за досить короткий проміжок часу. Але якщо у випадку наземних ліній зв'язку, за допомогою високошвидкісних оптичних кабельних мереж ця проблема має вирішення, то при бездротовому з'єднанні (як правило це радіоканал), пропускна швидкість такого каналу є досить обмеженою. Тобто при передачі зображень великих розмірів по радіоканалу, або іншій ситуації що обмежує швидкість передачі необхідно застосовувати певні методи та алгоритми кодування для більш оптимального використання ширини каналу зв'язку. Найбільш очевидною основою для зменшення кількості даних, що передаються є те що кінцевому користувачу в більшості випадків не потрібна повна деталізація, оскільки роздільна здатність сучасних моніторів, які використовуються в якості пристроїв відображення, рідко перевищує 2 мегапікселі. Якщо і необхідна висока деталізація, то лише невеликої частини зображення.

Для забезпечення передачі таких великих зображень по мережі Internet, де також не завжди є можливість передавати дані з великою швидкістю, фірмою Gigapan було розроблене програмне забезпечення, яке використовує алгоритм пірамідального розкладу. Принцип дії полягає у тому, що велике зображення розбивається на плитки, або тайли (від англ. tile — плитка). Чотири розташовані поряд тайли формують тайл вищого рівня з меншою у чотири рази роздільною здатністю. Таким чином з таких тайлів отримується оригінальне зображення з меншою роздільною здатністю. Процес повторюється з отриманим зображенням доти, поки отримане зображення не буде мати роздільну здатність оптимізовану під засіб відображення користувача. А сама структура даних не буде нагадувати піраміду. Для зберігання отриманих оброблених тайлів використовують стандартні алгоритми стиснення Jpeg, Jpeg2000.

Проблема полягає в тому, що отримана структура як правило займає на 50% більше місця ніж оригінальне зображення. І при передачі цих даних дані по суті передаються кінцевому користувачу стільки разів, стільки рівнів піраміди запросив користувач.

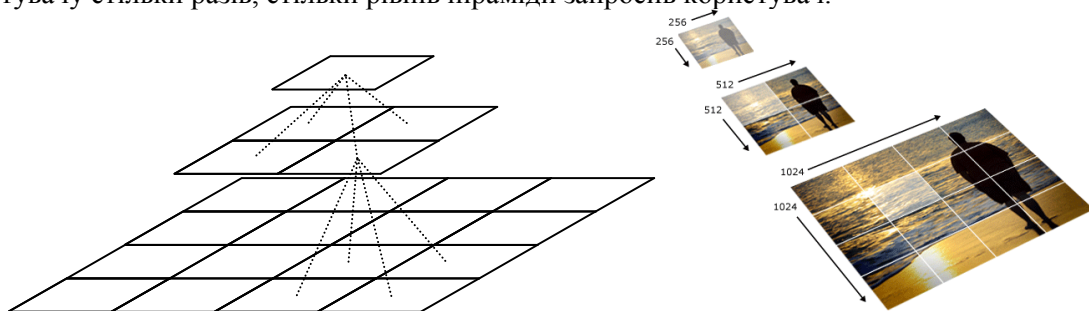


Рис. 4 Розбиття зображення.

Аналогічний алгоритм для зберігання та передачі великих зображень використовують у своїх розробках фірми Microsoft, технології DeepZoom [9], HD View [10], які в свою чергу базуються на технології SeaDragon [11]. Схожа технологія використовується у картографічних системах GoogleEarth, Яндекс карти, Wikimapia та інших.

Таким чином, можна сказати що не вирішеною залишається проблема оптимізації зберігання та передачі великих зображень. Тобто потрібна технологія роботи з великими

зображеннями, яка дозволить мінімізувати завантаження каналів зв'язку та уникнути дублювання у збереженні та передачі ЗВР.

Оскільки зображення по суті є двомірним сигналом, то доцільно спочатку розглянути рішення задачі для одномірного випадку. Нехай маємо цифровий сигнал $S = \{S_1, \dots, S_n\}$ з частотою дискретизації f_s , великого розміру ($n \gg 1000$). Оскільки оператор не в змозі сприймати цю інформацію повністю то передача цього сигналу S повністю (усі n значень) не має сенсу. Проте користувача цікавить структура сигналу S в цілому, а також можливість детально аналізувати окремі області (підмножини) S_i повного сигналу S ($S_i \in S$). В контексті поставленої задачі відомі алгоритми стиснення з втратами не здатні вирішити проблему мінімізації об'єму інформації, яка передається через канал зв'язку, оскільки повна деталізація на окремих фрагментах S_i неможлива через втрати на етапі стиснення. До того ж множини S_i апріорі невідомі, та задаються користувачем в процесі аналізу сигналу S в цілому, або в процесі деталізації.

Найпростішим узагальненим описом S , є його середнє значення \bar{S}_0 . Фактично \bar{S}_0 є числом, що презентуватиме сигнал S одним значенням. З точки зору фізичного змісту, \bar{S}_0 отримаємо розглядаючи сигнал S з такої відстані, коли його довжина відповідає одному пікселю. З статистичної точки зору \bar{S}_0 є оцінкою за МНК для моделі-константи. З точки зору цифрової фільтрації це значення пов'язане із низькочастотною фільтрацією не рекурсивним цифровим фільтром з N коефіцієнтами рівними $b_i = 1/N, i = \overline{1, N}$.

АЧХ такого фільтра:

$$H(\omega) = \frac{\sin(\omega \tau / 2)}{\omega \tau / 2}, \text{ де } \tau = \frac{N}{f_s};$$

Завдяки такій фільтрації частота дискретизації сигналу може бути знижена в N разів. Тобто, відлік \bar{S}_0 буде презентувати найнижчу частоту сигналу S (константу). Іншу частину сигналу (високочастотну) отримаємо, як вектор:

$$S_h = S - \bar{S} \quad \bar{S} = \{S_1 = \bar{S}_0, S_2 = \bar{S}_0, \dots, S_N = \bar{S}_0\}$$

Такій процедурі відповідає фільтрація фільтром верхніх частот з АЧХ $H_h(\omega)$. Фільтри $H(\omega)$ і $H_h(\omega)$ утворюють дзеркальну пару, яка дозволяє використовувати розклад і синтез сигналу S врахуванням зниження частоти дискретизації $\downarrow N$

Фільтр НЧ, що формує ${}^1S_{h_1}$ і ${}^2S_{h_1}$ має в k разів нижчий порядок, ніж $H(\omega)$ на попередньому етапі розкладу, та має АЧХ:

$$H_1(\omega) = \frac{\sin(\omega \tau / 4)}{\omega \tau / 4} = \frac{\sin(\omega \tau / 2k)}{\omega \tau / 2k}$$

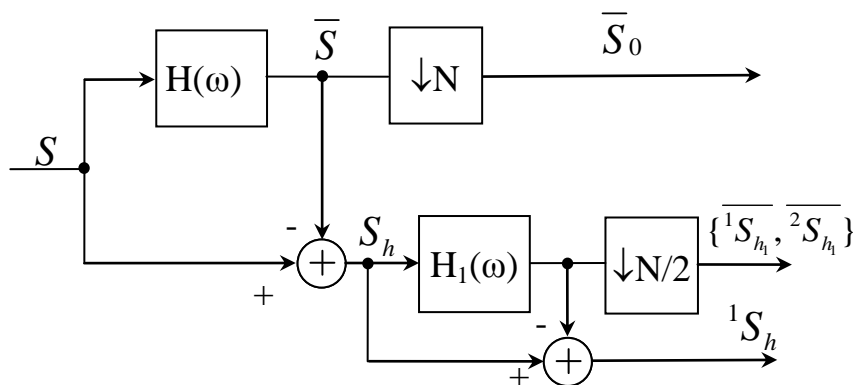


Рис. 5.

Тоді значення \bar{S}_0 презентуватиме \bar{S} в цілому, а S_h буде його деталізацією.

Показана на схемі (рис. 5) послідовність 2-х етапів розкладу можна застосувати до нескінченної послідовності та в режимі реального часу. Зрозуміло, що показану схему можна продовжити на необхідне число етапів розкладу.

Для 2-х вимірної задачі вирішення цієї задачі запропоновано алгоритм кодування зображень, що працює за принципом декомпозиції квадродерева [12,13]. Вузол ділить дерево на чверті. Вузол являє собою оригінальне зображення, а чверті це частини цього зображення (рис. 6)

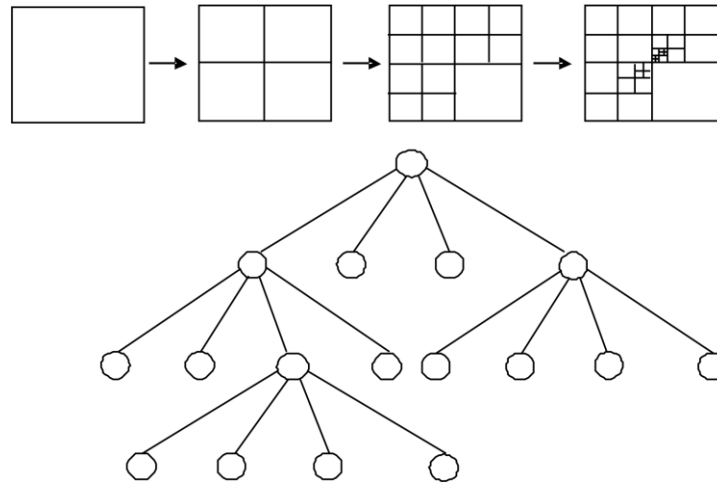


Рис. 6.

Такий алгоритм дуже добре реалізується як рекурсивний. Для кожної частини зображення викликається рекурсивна функція, яка розраховує середнє значення та віднімає його від цієї частини зображення, аналізує чи треба розбивати це зображення, якщо так, то знову ділить на чотири частини для яких викликається ця ж сама функція. Розраховане середнє значення зберігається в вузлах дерева. Глибина поділу на гетерогенних частинах зображення може досягати одного пікселя. Гомогенні частини не розбиваються.

Оскільки використання такого алгоритму для кодування великих зображень потребує значних апаратних ресурсів, то доречно використовувати для цих цілей сучасні відеокарти, наприклад архітектури Fermi, де встановлено кілька десятків розрахункових ядер, що працюють паралельно.

Можливість передачі частини зображення з необхідною роздільною здатністю здійснюється за рахунок обходу дерева по ширині до певної глибини. Поєднання декомпозиції зображення квадродерева з інтерполяцією вузлів-зображень дозволяє отримати гібридний алгоритм стиснення зображення, який може давати користувачу можливість отримувати зображення з необхідним йому співвідношенням якості (деталізація) – розмір (час передачі по лініям зв'язку).

При побудові рекурсивних алгоритмів декомпозиції зображень за допомогою квадродерева та використання фільтрів низьких частот, при збільшенні зображень значно зростає ресурсоемність задачі, потреби в оперативній пам'яті, так і в кількості процесорного часу. Таким чином виникає логічне питання адаптації такого алгоритму для роботи на багатоядерних обчислювальних пристроях, і яких найбільш поширеними є сучасні графічні процесори (Наприклад архітектури Fermi, фірми Nvidia)

Такий підхід дозволяє передавати спочатку зображення з найменшою деталізацією, виділене фільтром низьких частот, окремим випадком якого є середнє значення. Розвитком такого алгоритму є застосування в якості фільтрів низьких частот двомірні сплайн функції та двомірний LSS розклад [14]. Також такий підхід можна реалізувати для сигналів з іншою розмірністю. Наприклад для кодування відео застосувати розбиття октодерева. Тобто для того щоб обробити послідовність кадрів вводиться третій вимір – час. При застосуванні такого алгоритму можливо реалізувати формат передачі відео у якому бітрейт потоку буде автоматично підстроюватись під можливості каналу користувача.

Впровадження таких алгоритмів кодування великих зображень дозволить використовувати всі технічні можливості, які дають камери з великою роздільною здатністю особливо при їх використанні у сфері безпілотних літаючих апаратів, а саме зникне необхідність чекати доки до місця знаходження оператора буде доставлений носій даних з відзнятим матеріалом, оператор буде мати можливість ознайомлюватись з даними отриманими камерою в режимі онлайн, без підвищених вимог до пропускної спроможності наявного каналу зв'язку, що відповідно з одного

боку надає можливість спростити прийомопередавальну апаратуру, а з іншого підвищити швидкість прийняття рішень на рахунок отриманої інформації.

- 1 Белова И. Б., Китаев В. М. Цифровые технологии получения рентгеновского изображения - принципы формирования и типы (обзор литературы)//Медицинская визуализация.-2000-№1-С.33-40
- 2 Рекламний буклет фірми GE Healthcare, Лидер по разрешающей способности изображения среди компьютерных томографов. Система Discovery™ CT750 HD, 26 с.
- 3 Сечин А.Ю. Эпоха цифровой аэросъемки. //Пространственные данные. – 2009. – №3 – с. 29-35.
- 4 О.Н. Зинченко, А.Б. Елизаров Цифровые камеры для аэрофотосъемки. Обзор моделей (декабрь 2010), Ракурс, Москва, Россия. 2010
- 5 Електронний ресурс/ опис обладнання фірми Gigapan Режим доступу:<http://gigapan.com/cms/shop/epic-100>
- 6 http://en.wikipedia.org/wiki/Largest_photographs_in_the_world
- 7 <http://en.wikipedia.org/wiki/ARGUS-IS>.
- 8 Brady, D. J., *et al.* *Nature* **486**, 386–389 (2012)., Електронний ресурс/<http://www.nature.com/nature/journal/v486/n7403/full/nature11150.html>)
- 9 Електронний ресурс/ опис технології DeepZoom Режим доступу: [http://msdn.microsoft.com/en-us/library/cc645050\(v=vs.95\).aspx](http://msdn.microsoft.com/en-us/library/cc645050(v=vs.95).aspx)
- 10 Електронний ресурс/ опис технології HD View Режим доступу:<http://hdviewsl.codeplex.com/>
- 11 Електронний ресурс/ опис технології SeaDragon Режим доступу: <http://www.asp.net/ajaxlibrary/AjaxControlToolkitSampleSite/Seadragon/Seadragon.aspx>
- 12 **Павлидис Т.** Алгоритмы машинной графики и обработки Изображений. / Павлидис Т. // Вісник Національного університету “Львівська політехніка”. – Львів, Видавництво Мир, 1986 – 396с
- 13 **Yi-Chen Tsai, Ming-Sui Lee, Meiyin Shen, and C.-C. Jay Kuo.** A quad-tree decomposition approach to cartoon image compression. IEEE 8th Workshop on Multimedia Signal Processing, 17(6):456–460, October 2006.
- 14 **І.В.Шелевицький, М.О.Шутко, В.М.Шутко, О.О. Колганова** - Слайни в цифровій обробці даних і сигналів. – Кривий Ріг, Видавничій Дім 2007р. –232., іл.