

**К.т.н, доцент Маслянюк П.П, магістрант Михальський Ю.Л.**

**Національний технічний університет України  
«Київський політехнічний інститут»**

## **АНАЛІЗ МЕТОДІВ ТА РОЗРОБКА МЕТАМОДЕЛІ РОЗШИРЕНОЇ РЕАЛЬНОСТІ**

### **Abstract**

*Pavlo P. Maslyanko, assoc. prof., PhD; Yuriy Mykhalskiy, student  
Methods analysis and creation metamodel of augmented reality*

*Abstract deals with metamodel of augmented reality, which is based under new search feature points search algorithm. Goal of metamodel is purpose to use any invariant images. It combines positive aspects of two common algorithms: SURF and SIFT. Main idea is using Gaussian pyramid to correct feature points data from SURF.*

### **Вступ**

Набувають використання комп'ютерні тривимірні моделі для візуалізації об'єктів. Перспективні розробки надають можливість доповнювати (розширювати) зображення, що отримані з допомогою відеокамер чи веб-камер [1].

Задача доповнення зображення нетривіальна і вимагає створення метамоделі розширеної реальності. Обробка зображень отриманих з відеопотоку за допомогою спеціальних алгоритмів дозволяє отримати так звані особливі точки (feature points). У результаті отримують матрицю переходу від одного зображення (наприклад, еталонного) до поточного, після чого формується тривимірна модель, яка оброблюється та доповнюється необхідними елементами [2,3].

### **Термінологія**

*Розширена реальність* - доповнення реальності будь-якими віртуальними елементами [1].

*Блоб* — регіон, яскравість якого сильно відрізняється від сусідніх регіонів [1].

*Особлива точка* — точка в образі, яка має чітко визначені позицію і може бути однозначно виявлена [2].

## Постановка задачі

Об'єкт дослідження – системи розширеної реальності

Предмет дослідження – доповнення реального зображення з відеопотоку інтерактивними елементами.

Метою роботи є аналіз методів та розробка метамоделі розширеної реальності.

## Огляд існуючих рішень

В основі розширеної реальності лежать різноманітні методи комп'ютерного зору, в основному, пов'язані з відеоспостереження (video tracking).

Найбільш поширеними методами розпізнавання зображень є [1]:

1. збудовані на кореляції,
2. метод Харріса-Стівенса,
3. SIFT (Scale-invariant feature transform),
4. SURF (Speeded Up Robust Features).

З методів проєктивної геометрії найбільшого розповсюдження знайшли методи засновані на гомографії.

## Метамоделі розширеної реальності

На засадах цих алгоритмів базується запропонована метамоделі.

Метамоделі можна зобразити наступною діаграмою у нотації UML:

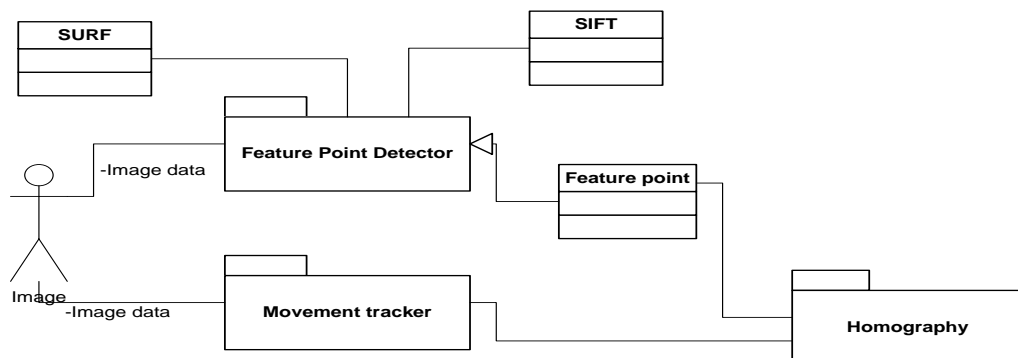


Рис 1. Метамоделі розширеної реальності представлена у нотації UML

*Компонент Movement tracker.* Використовується у випадку погані визначеності особливих точок, за поріг застосування використовується загальна кількість особливих точок.

*Компонент Feature point detector.* Він має 2 конкретні реалізації - пакети SURF та SIFT.

За основу оберемо алгоритм SURF та покращимо пошук особливих точок за допомогою іншого метода. Для цього збудуємо карту зображення отриману з SURF:

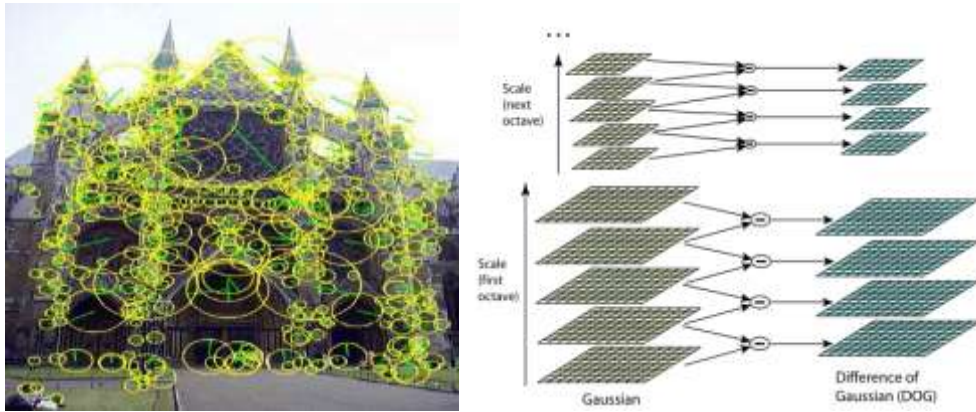


Рис 2. Карта особливих точок отриманих за алгоритмом SURF(ліворуч) та піраміда гаусіанів отримана з допомогою алгоритму SIFT(праворуч)

Після чого, перевіримо це зображення з допомогою алгоритма SIFT. Це дозволить скористатися інтегральними можливостями другого алгоритму та перевагами інваріантності першого. Координати особливих точок з субпіксельною точністю можна отримати із (1)

$$D \begin{matrix} \bar{x} \\ \bar{y} \end{matrix} = D + \frac{\partial D^T}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} x \quad (1)$$

Тут  $D$  - функція DoG,  $X = \begin{matrix} \bar{x} \\ \bar{y} \end{matrix} = \begin{matrix} x, y, \sigma \end{matrix}$  - вектор зміщення відносно точки розкладання, перша похідна DoG - градієнт, друга похідна DoG - матриця Гессе. Екстремум многочлена Тейлора відшукується шляхом обчислення похідної і порівнювання її до нуля. У результаті отримаємо зміщення точки обчисленого екстремуму, щодо точного

$$\hat{x} = -\frac{\partial^2 D^{-1}}{\partial x^2} \frac{\partial D}{\partial x}$$

*Компонент Homography.*

Є відповідальним за пошук гомографії між двома наборами точок векторами  $\bar{x} = \begin{matrix} x_1, x_2, \dots \end{matrix}$  та  $\bar{y} = \begin{matrix} y_1, y_2, \dots \end{matrix}$  - точками на вихідному зображенні та поточному. Для пошуку гомографії скористаємось методом — Лівенберг-Маркварда. Він дозволить мінімізувати наступне відношення:

$$S \cdot H = \sum_{i=1}^n y_i - H \cdot x_i$$

З цього рівняння можна буде отримати матрицю гомографії  $H$ , яку можна використовувати у подальшій обробці.

## Результати

Для перевірки метамоделі, була використана стандартна бібліотека зображень [4]. З неї було обрано частину, яка містила зображення різних будівель (розмір вибірки – 750 зображень). Для стандартних алгоритмів були отримані наступні результати:

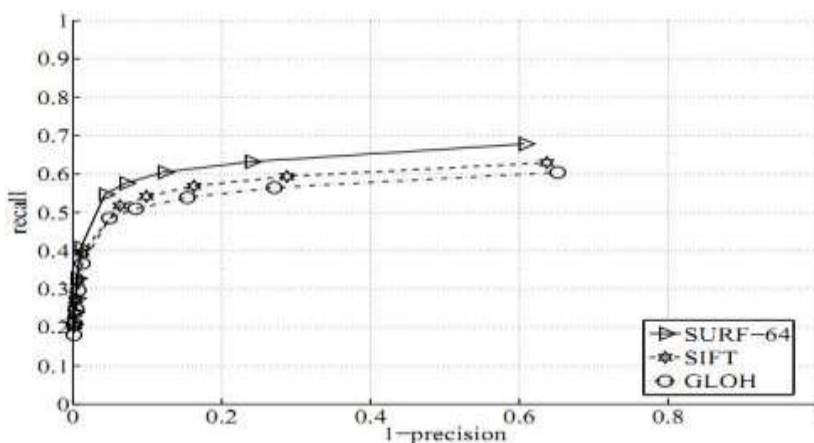


Рис 3. Точність пошуку особливих точок за алгоритмами SURF-64, SIFT, GLOH

З рисунку видно, що модель SURF надає найбільшу точність пошуку, з усіх наведених.

У подальшому, для тестування, була використана збудована модель. В ній використовувалась піраміда гаусіанів у кількості – 8 та розмірність дескрипторів SURF – 32. Проводилось порівняння із алгоритмами класу SURF з різною розмірністю дескрипторів.

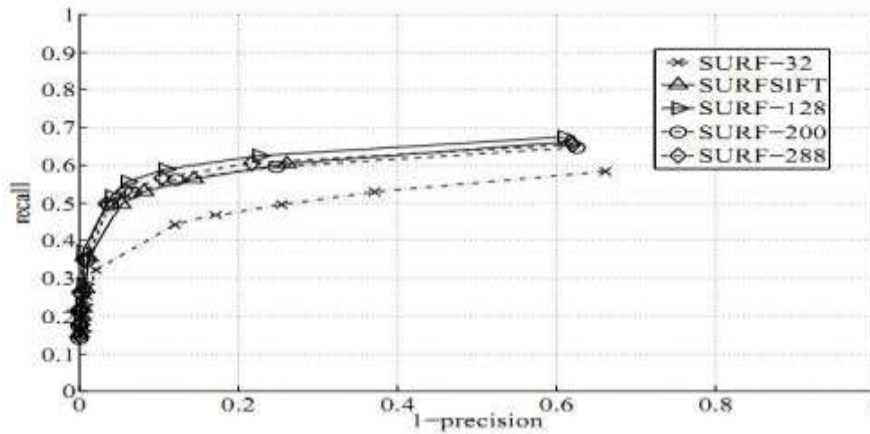


Рис 4. Порівняння алгоритмів класу SURF із запропонованим SURFSIFT

З отриманих даних видно, що збудований алгоритм програє алгоритму з більшою розмірністю дескриптору – SURF-128, але швидкодія збудованого алгоритму є на рівні SURF-32.

### Висновки

Запропонована метамодель дозволяє скористатись особливими точками навіть у випадку поганої визначеності цих точок, що досягається за рахунок доповнення результатів одного алгоритму результатами іншого. Це використовується за основу при обчисленні гомографії для пошуку матриці переходу у двовимірному просторі чи тензора у тривимірному.

Подальша робота у даному напрямку дозволить оптимізувати процес побудови січних площин та використання піраміди гауссіанів.

### Література

1. *Hartley R., Zisserman A. Multiple View Geometry in Computer Vision.* – Cambridge University Press, 2001. – pp. 214-300, 458.
2. *Shapiro, Linda and George C. Stockman. Computer Vision.*, Prentice Books, Upper Saddle River, 2001. - pp. 134,257.
3. *C. Harris and M. Stephens. "A combined corner and edge detector".* Proceedings of the 4th Alvey Vision Conference. pp. 147–151. <http://www.bmva.org/bmvc/1988/avc-88-023.pdf>.
4. <http://www.cs.cmu.edu/~cil/v-images.html>