

*Драган Д.Д., Поляков А.Ю. — рецензент Петренко А.І.
Інститут прикладного системного аналізу НТУУ «КПІ», Київ, Україна*

Методи виявлення і розпізнавання рухомих об'єктів при відеоспостереженні

На сьогоднішній день, комп'ютерний зір (Computer Vision) став невід'ємною частиною науки і технології виготовлення машин, які бачать. Це пов'язано з розробкою та впровадженням алгоритмів, які можуть автоматично обробляти візуальні дані, для розпізнавання об'єктів, їх відстеження та відновлення їх форми і просторового розташування. Однією із фундаментальних проблем комп'ютерного зору є «розуміння зображень» чи «семантична інтерпретація», тобто інтерпретація сцени, що містяться в зображенні, як сукупності значущих сутностей. Це може включати розбір інформації в сцені на різних рівнях.

В якості попередньої обробки відеопослідовності була використана ефективна ієрархічна фільтрація (Hierarchical Filter) [1], яка забезпечила зменшення сторонніх рухів із фону сцени. На відміну від використання методу Spatial-Temporal Interest Point (STIPs) [2] було вирішено, що просторова і часова інформація має розглядатися окремо. Часова інформація отримується шляхом використання методу Motion Histogram Image (МНІ) [3], в основі якого лежить віднімання кадрів, а для отримання просторової - двовимірний метод Харіса для кутів (2-D Harris corners) [4]. Використання МНІ дозволяє уникнути неправильного відстеження точок у відеопослідовності. МНІ-пікселі з яскравішою інтенсивністю являють собою рухомі об'єкти з більш пізнім рухом, тому вони використовуються для створення шаблонів. Тоді отримані шаблони руху і отримані із 2-D Harris кути використовуються для виявлення точок інтересу.

Розглядаються тільки кути з найостаннішого руху як точки інтересу. Так як ізольований напрямок руху пікселів часто є небажаним рухом або шумом в порівнянні з сусідніми пікселями, було розроблено алгоритм для усунення цього небажаного руху. По-перше, використати згладжувальний фільтр для глобального просторового руху до градієнтів МНІ. Потім в кожній точці інтересу, обчислюється структурна близькість між будь-яким пікселем в локальному регіоні і точкою інтересу локальної рухомої зони. Таким чином, можна підсилити або послабити рух до пікселя в залежності від його структурної близькості до точки інтересу. Для отримання часових характеристик, був використаний новий часовий дескриптор гістограм орієнтованих градієнтів в МНІ (HOG-МНІ). Просторові характеристики отримуються за рахунок використання методу HOG [5] до інтенсивності зображення. Отримані характеристики векторів, які містять як HOG (просторові характеристики), так і HOG-МНІ (часові характеристики) використовуються для розпізнавання, на основі класифікатора Gaussian mixture model (GMM) [6].

Література. 1. Y. Tian, L. Cao, Z. Liu, and Z. Zhang, Hierarchical Filtered Motion for Action Recognition in Crowded Videos, IEEE Transactions on Systems, Man, and Cybernetics–Part C: Applications and Reviews, 2011. <http://media-lab.engr.ccnycunyu.edu/Paper/2011/SMCC-HumanAction.pdf> 2. L. Cao, Y. Tian, Z. Liu, B. Yao, Z. Zhang, and T. Huang, Action Detection using Multiple Spatial-Temporal Interest Point Features, IEEE International Conference on Multimedia & Expo (ICME), 2010. <http://media-lab.engr.ccnycunyu.edu/Paper/2010/ICME2010.pdf> 3. James W. Davis, Recognizing Movement using Motion Histograms, MIT Media Laboratory Perceptual Computing Section Technical Report №487, <http://hd.media.mit.edu/tech-reports/TR-487.pdf> 4. 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> 5. N. Dalal and B. Triggs, Histograms of Oriented Gradients for Human Detection, INRIA Rhone-Alps, 655 avenue de l'Europe, Montbonnot 38334, France. <http://lear.inrialpes.fr/people/triggs/pubs/Dalal-cvpr05.pdf> 6. Douglas Reynolds, Gaussian Mixture Models, MIT Lincoln Laboratory, 244 Wood St., Lexington, MA 02140, USA. http://www.ll.mit.edu/mission/communications/ist/publications/0802_Reynolds_Biometrics-GMM.pdf