

Преобразование инфракрасных изображений в видимый диапазон для визуализации на индикаторе пилота

В.В. Князь^{1,2}, М.И. Козырев^{1,3}, А.Н. Бордодымов¹
vl.kniaz@gosniias.ru | j18r11@gmail.com | bordodymov@gmail.com

¹ФГУП «ГосНИИАС», 125319, Москва, ул. Викторенко, 7;

²Московский физико-технический институт (государственный университет), Москва, Россия;

³Московский Государственный Технический Университет Имени Н. Э. Баумана, Москва, Россия;

Камеры дальнего инфракрасного диапазона позволяют повысить ситуационную осведомлённость экипажа в условиях ограниченной видимости и в ночное время. Подобные камеры устанавливаются на современные гражданские самолёты в составе системы улучшенного видения. При этом интерпретация тепловизионного изображения экипажем для верной оценки текущей ситуации требует определённого опыта в силу того, что вид сцены в инфракрасном диапазоне сильно отличается от видимого диапазона и подвержен изменениям при смене времени суток и времён года. В данной работе рассматривается глубокая генеративно-сопоставительная нейронная сеть для автоматического преобразования тепловизионных изображений в семантически подобные цветные изображения видимого диапазона.

Ключевые слова: визуализация, глубокие свёрточные нейронные сети, индикатор пилота, визуальная аналитика.

Thermal-to-color image translation for visualization on the pilot's display

V.V. Kniaz^{1,2}, M.I. Kozyrev^{1,3}, A.N. Bordodymov¹
vl.kniaz@gosniias.ru | j18r11@gmail.com | bordodymov@gmail.com

¹State Res. Institute of Aviation Systems (GosNIIAS), 125319, 7, Victorenko str., Moscow, Russia;

²Moscow Institute of Physics and Technology (MIPT), Russia;

³Bauman Moscow State Technical University (BMSTU), Moscow, Russia

Long range infrared cameras may provide increasing crew situational awareness in limited vision and night conditions. Similar cameras are installed in modern civil aircrafts as part of an improved vision system. Correct thermal image interpretation by the crew requires certain experience, due to the fact that view of the scene very different from the visible range and may change within time of day and season. This paper discusses the deep generative-adversary neural network to automatically convert thermal images to semantically similar color images of the visible range.

Keywords: visualization, deep convolutional neural networks, pilot primary display, visual analytics.

1. Введение

Повышение ситуационной осведомлённости экипажа является залогом обеспечения безопасности полёта. В настоящее время в состав бортового оборудования многих гражданских самолётов входят системы улучшенного видения. Данная система включает датчик инфракрасного диапазона, который производит обзор кабиной обстановки в передней полусфере, и вычислитель который принимает видеосигнал и выводит его на индикатор пилота. Тепловизионный инфракрасный датчик обеспечивает отображение видимых объектов и рельефа местности в условиях низкой освещённости и затруднённой видимости.

К числу недостатков системы улучшенного видения с инфракрасным датчиком можно отнести сложности при интерпретации тепловизионного изображения. В силу того, что тепловое излучение объектов в разную погоду может быть различным, то их вид на кадре системы улучшенного видения может сильно варьироваться от времени года и времени суток. Например, взлётно-посадочная полоса может быть светлой на тёмном фоне в солнечный день и тёмной на светлом во время дождя. Для облегчения задачи обнаружения пилотом визуальных ориентиров представляется целе-

сообразным предварительная обработка кадра системы улучшенного видения для преобразования изображения инфракрасного диапазона в видимый диапазон.

В данной работе рассматривается метод преобразования монохромных тепловизионных изображений в цветные изображения видимого диапазона. Метод использует модифицированный вариант генеративно-сопоставительной сети ColorMatchGAN. Представлена архитектура сети. Для обучения и тестирования сети собрана обучающая выборка с помощью квадрокоптера DJI Mavic PRO, оснащённого камерами видимого и дальнего инфракрасного диапазона. Представлена методика полуавтоматического геометрического совмещения кадров видимого и инфракрасного диапазонов. Модифицированная сеть ColorMatchGAN обучена на собранной выборке. Тестирование производилось на независимой выборке объёмом в 400 кадров

2. Анализ работ в данной области

Системы повышения ситуационной осведомлённости на основе машинного зрения получили широкое распространение в течение последнего десятилетия [2, 11]. Наиболее широкое распространение получили системы на основе инфракрасных датчиков дальнего ин-

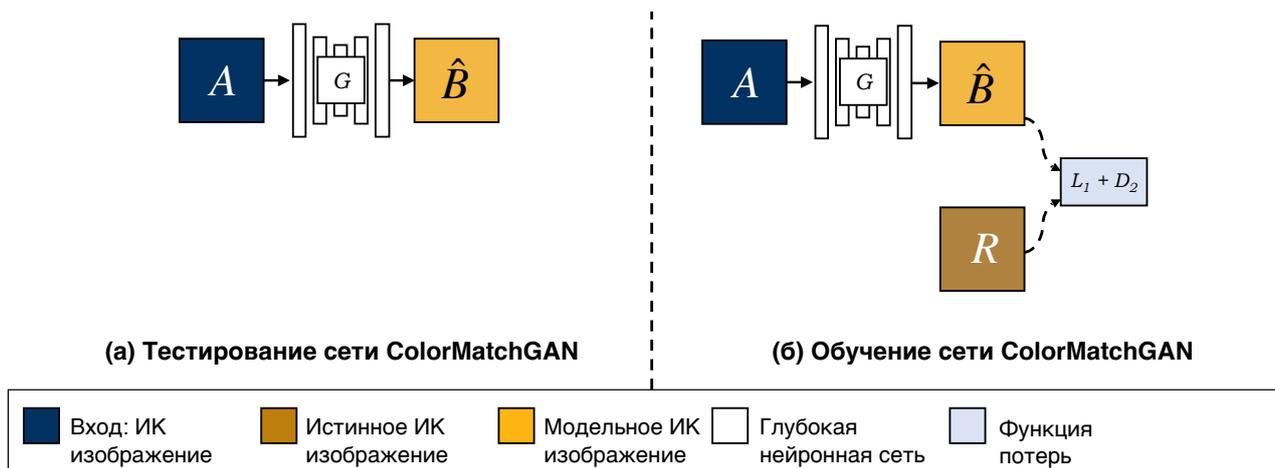


Рис. 1. Архитектура сети ColorMatchGAN.

фрактального диапазона (8-14 мкм), которые производят обзор кабины обстановки в направлении движения воздушного судна [2, 11]. Такие системы принято называть системами улучшенного зрения (СУВ). Основными критериями качества систем улучшенного зрения является дальность обнаружения взлётно-посадочной полосы и препятствий на её поверхности или в воздухе. Для повышения дальности обнаружения объектов предложены разнообразные алгоритмы повышения качества изображения [7].

Несмотря на существенное повышение ситуационной осведомлённости, обеспечиваемое современными системами улучшенного зрения, интерпретация тепловизионных изображений может вызывать у экипажа значительные сложности. Целесообразной является предобработка тепловизионного изображения, производящая предсказание цветов объектового состава и фона для облегчения интерпретации наблюдаемой сцены. В течение последних пяти лет активно развиваются нейросетевые методы обработки изображений на основе генеративно-сопоставительных нейронных сетей [3, 4]. Основная идея генеративно-сопоставительного подхода заключается в обучении двух соревнующихся сетей: генератора G и дискриминатора D . Цель генератора выучить заданное распределение изображений $\mathcal{B} \subset \mathbb{R}^{W \times H \times C}$ и научиться воспроизводить его на основе вектора шума z или входного изображения A . Целью дискриминатора является бинарная классификация входного изображения на классы: «реальное» и «модельное». «Реальные» изображения $B \in \mathcal{B}$ принадлежат пространству выходных изображения. «Модельные» изображения \hat{B} являются результатом работы сети генератора G . Сопоставительная функция потерь накладывает на сеть генератор штраф, в случае, если сеть дискриминатор производит верную классификацию изображений \hat{B} классом «модельное». Таким образом, сеть генератор пытается строить максимально правдоподобные изображения \hat{B} , чтобы запутать сеть дискриминатор.

В последние годы был предложен ряд работ по преобразованию спектрального диапазона изображений на основе генеративно-сопоставительных нейросетей [1, 5, 10]. В данной работе рассматривается модификация архитектуры ColorMatchGAN [5], предназначенной для предсказания цветных изображений по тепловизионным изображениям.

3. Метод преобразования инфракрасных изображений в видимый диапазон

Целью рассматриваемого метода является отображение входного изображения $A \in \mathbb{R}^{W \times H}$ дальнего инфракрасного диапазона в цветное изображение видимого диапазона $B \in \mathbb{R}^{W \times H \times 3}$. Требуемое отображение $G : A \rightarrow \hat{B}$ реализуется с помощью модифицированной сети генератора на основе архитектуры ColorMatchGAN [5]. В данном разделе рассмотрены условные генеративно-сопоставительные нейронные сети, лежащие в основе разрабатываемого метода, представлена модифицированная архитектура сети и методика подготовки обучающей выборки.

Архитектура сети

Генеративно-сопоставительные сети используют [3] сопоставительную функцию потерь для снижения вероятности переобучения сети. Генеративно-сопоставительные сети создают изображение \hat{B} для заданного вектора случайного шума z , $G : z \rightarrow \hat{B}$ [3, 4]. Условные генеративно-сопоставительные сети получают дополнительную информацию A в дополнение к вектору z , $G : \{A, z\} \rightarrow \hat{B}$. Обычно, A – это изображение, которое трансформируется генеративной моделью G . Дискриминативная модель тренирована для того чтобы различать «реальные» изображения из целевого домена B от «модельных» \hat{B} , созданных генератором. Каждая модель тренируется одновременно. Дискриминативная модель создает сопоставительную потерю, которая заставляет генератор производить «модельные» \hat{B} изображения, которые не отличить от «реальных» B .

Архитектура сети ColorMatchGAN [5] включает генератор U-Net [8] и дискриминаторе PatchGAN [4]. Архитектура сети ColorMatchGAN представлена на рисунке 1.

Построение обучающей выборки

Для обучения модифицированной архитектуры сети использовалась обучающая выборка LAERT [6]. Данная выборка собрана с помощью квадрокоптера DJI Mavic PRO, оснащённого встроенной камерой видимого диапазона, и дополнительной камерой дальнего инфракрасного диапазона (8-14 мкм) MH-SM576-6 с разрешением 640×480 пиксел. Общий вид квадрокоптера приведён на рисунке 2.



Рис. 2. Вид квадрокоптера Mavic PRO с камерами видимого и инфракрасного диапазона.

В силу того, что камера видимого диапазона закреплена на гиросtabilизированном подвесе, а тепловизионная камера жёстко соединена с корпусом, между цветными и тепловизионными изображениями присутствует динамическое расхождение. Для устранения геометрического расхождения разработана методика полуавтоматического совмещения изображений двух диапазонов.

Методика совмещения изображений двух диапазонов основана на использовании матрицы гомографии.

$$H = \begin{bmatrix} h_{00} & h_{01} & h_{02} \\ h_{10} & h_{11} & h_{12} \\ h_{20} & h_{21} & h_{22} \end{bmatrix}$$

Пусть (x_v, y_v) – точка на изображении видимого диапазона и (x_t, y_t) – точка на тепловизионном изображении, в том же физическом месте. Тогда гомография H связывает их следующим образом

$$H = \begin{bmatrix} x_v \\ y_v \\ 1 \end{bmatrix} = H \begin{bmatrix} x_t \\ y_t \\ 1 \end{bmatrix} = \begin{bmatrix} h_{00} & h_{01} & h_{02} \\ h_{10} & h_{11} & h_{12} \\ h_{20} & h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} x_t \\ y_t \\ 1 \end{bmatrix}$$

Если параметры матрицы гомографии известны, то можно найти переход от заданной точки на снимке в видимом диапазоне (x_v, y_v) к соответствующей точке (x_t, y_t) на изображении инфракрасного диапазона. Для расчёта матрицы гомографии необходимо знать,

как минимум, четыре соответствующие точки на двух изображениях.

Очевидно, что процесс автоматической расстановки пар точек на всех кадрах видеопоследовательности трудоёмкий процесс. Для автоматизации задачи предлагается использовать прослеживание точек между кадрами с помощью кросс-корреляции. Четыре соответствующие точки расставляются на первом кадре видеопоследовательности и прослеживаются до того момента, пока они видны в поле зрения камеры. Координаты (x_v^i, y_v^i) , (x_t^i, y_t^i) каждой точки, на каждом кадре i , заносятся в массив. После этого для каждого элемента полученного массива производится преобразование кадра видимого диапазона к кадру инфракрасного диапазона.

Предложенная методика была реализована в виде сценария на языке Python. Для прослеживания соответствующих точек использовалось API пакета трёхмерного моделирования Blender. Примеры из обучающей выборки приведены на рисунке 3.

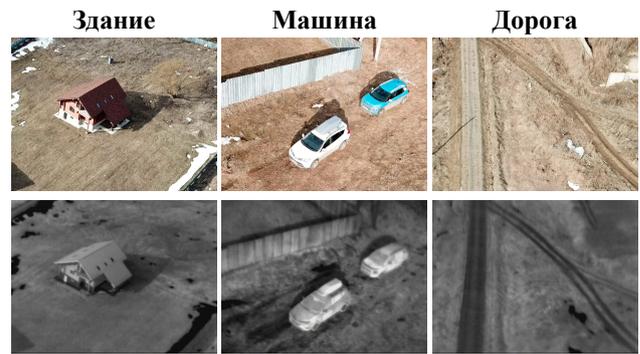


Рис. 3. Примеры из обучающей выборки LAERT.

4. Экспериментальное тестирование метода

ColorMatchGAN обучался на независимом тестовом разделе обучающей выборки LAERT, с использованием библиотеки PyTorch. При обучении использовался графический процессор NVIDIA 1080 Ti. Процесс обучения занял 76 часов для генератора G и дискриминатора D . Для оптимизации сети использовался алгоритм градиентного спуска Adam с начальной скоростью обучения 0.0002 и параметрами моментов $\beta_1 = 0.5$, $\beta_2 = 0.999$, аналогично работе [4].

Результаты экспериментального тестирования сети приведены на рисунке 4. Качественное сравнение результатов показывает, что сеть ColorMatchGAN обеспечивает прирост качества предсказываемых цветных изображений. Количественное тестирование по метрике LPIPS [9] показывает, что расстояние между истинными цветными изображениями и предсказанием ColorMatchGAN меньше аналогичного расстояния для изображений, предсказанных нейросетью pix2pix на 20%.

5. Заключение

Рассмотрен метод преобразования изображений дальнего инфракрасного диапазона в цветные изображения видимого диапазона. Предложенный метод основан на генеративно-сопоставительных нейронных сетях. Разработана и реализована в виде сценария на языке Python для библиотеки PyTorch модификация архитектуры сети ColorMatchGAN. Предложенная модификация состоит в переходе в цветовое пространство Lab для повышения равномерной сходимости процесса обучения. Произведена обработка многоспектральной обучающей выборки LAERT для синхронизации и геометрического совмещения кадров видимого и инфракрасного диапазона. Сформированы обучающая выборка, объёмом 4000 кадров и независимая тестовая выборка объёмом 400 кадров.

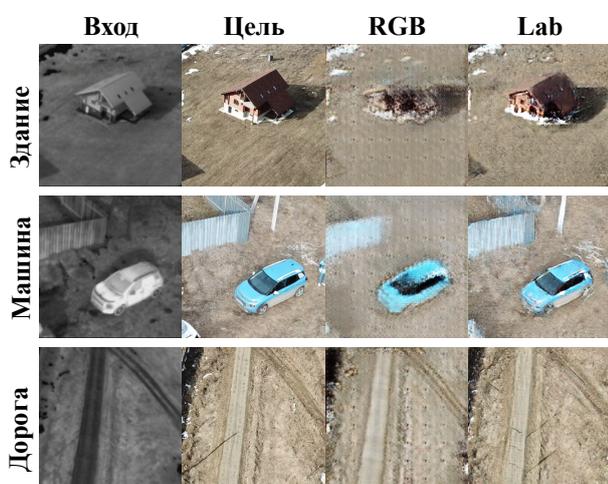


Рис. 4. Результаты экспериментального тестирования сети.

Произведено обучение нейросети ColorMatchGAN и сети аналога pix2pix. Тестирование обученных сетей показало превосходство предложенной сети ColorMatchGAN в части реалистичности и интерпретируемости синтезируемых изображений. Анализ работы предложенного метода на видеопоследовательностях подтверждает работоспособность предложенного подхода.

6. Благодарности

Работа выполнена при поддержке Российского фонда фундаментальных исследований, грант РФФИ № 17-29-04410.

7. Литература

[1] Berg Amanda, Ahlberg Jorgen, Felsberg Michael. Generating Visible Spectrum Images From Thermal Infrared // The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops. — 2018. — June.

- [2] Arthur Jarvis J., Norman R. Michael, Kramer Lynda J. et al. Enhanced vision flight deck technology for commercial aircraft low-visibility surface operations. — 2013. — Access mode: <https://doi.org/10.1117/12.2016386>.
- [3] Generative adversarial nets / Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza et al. // Advances in neural information processing systems. — 2014. — P. 2672–2680.
- [4] Image-to-Image Translation with Conditional Adversarial Networks / Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A Efros // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). — IEEE, 2017. — P. 5967–5976.
- [5] Kniiaz V. V., Bordodymov A. N. LONG WAVE INFRARED IMAGE COLORIZATION FOR PERSON RE-IDENTIFICATION // ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. — 2019. — Vol. XLII-2/W12. — P. 111–116. — Access mode: <https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLII-2-W12/111/2019/>.
- [6] Knyaz Vladimir. Multimodal data fusion for object recognition. — Vol. 110590. — 2019. — P. 110590P. — Access mode: <https://doi.org/10.1117/12.2526067>.
- [7] Petro Ana Belén, Sbert Catalina, Morel Jean-Michel. Multiscale retinex // Image Processing On Line. — 2014. — P. 71–88.
- [8] Ronneberger Olaf, Fischer Philipp, Brox Thomas. U-net: Convolutional networks for biomedical image segmentation // International Conference on Medical image computing and computer-assisted intervention / Springer. — 2015. — P. 234–241.
- [9] The Unreasonable Effectiveness of Deep Features as a Perceptual Metric / Richard Zhang, Phillip Isola, Alexei A Efros et al. // The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). — 2018. — Jun.
- [10] ThermalGAN: Multimodal Color-to-Thermal Image Translation for Person Re-Identification in Multispectral Dataset / Vladimir V. Kniiaz, Vladimir A. Knyaz, Jiří Hladůvka et al. // Computer Vision – ECCV 2018 Workshops. — Springer International Publishing, 2018.
- [11] Vygolov Oleg, Zheltov Sergey. Enhanced, synthetic and combined vision technologies for civil aviation // Computer Vision in Control Systems-2. — Springer, 2015. — P. 201–230.

Об авторах

Князь Владимир Владимирович, начальник группы ГосНИИАС, аспирант ГосНИИАС. Его e-mail vl.kniiaz@gosniias.ru.

Козырев Марк Игоревич, техник ГосНИИАС. Его e-mail J18r1L@gmail.com.

Бордодымов Артём Николаевич, инженер ГосНИИАС. Его e-mail bordodymov@gmail.com.