

Обнаружение смазанных кадров

А.А. Саенко, А.С. Миронов

Тихоокеанский государственный университет, Хабаровск

Аннотация: Смазанные кадры представляют значительную проблему в различных областях, таких, как видеонаблюдение, медицинская визуализация и аэрофотосъемка, при решении следующих проблем необходимы обнаружение и идентификация объектов, диагностика заболеваний на основе изображений, а также анализ и обработка данных с дронов для создания карт и проведения мониторинга. В данной статье предлагается метод обнаружения смазанных кадров с использованием нейросетевой модели. Принцип работы модели заключается в анализе изображений, представленных в частотной области в пространстве Хафа. Полученные результаты показывают, что предложенный метод обладает высокой точностью в обнаружении смазанных кадров и может быть использован в системах, где требуется высокая точность и четкость визуальных данных для принятия решений.

Ключевые слова: смазанные кадры, motion blur, blur, преобразование Хафа, спектральный анализ.

Введение

В современном мире обработка изображений и видео играет ключевую роль в различных областях, таких, как видеонаблюдение, медицина, строительство, архитектура и аэрофотосъемка [1]. В этих сферах качество видео напрямую влияет на точность и эффективность выполняемых задач. Например, в системах видеонаблюдения низкое качество изображения может препятствовать идентификации подозреваемых лиц и объектов [2, 3]. В медицинской диагностике нечеткие изображения могут привести к неправильной интерпретации данных и ошибкам в постановке диагноза.

Одной из распространенных проблем в обработке видео является обнаружение смазанных кадров, которые могут снижать качество анализа и восприятия информации [4, 5]. Смазанные кадры в видео являются результатом быстрого движения камеры или объекта, недостаточной экспозиции, неправильной фокусировки, и т.д., приводящих к размытому изображению. Это особенно критично в задачах, требующих высокой точности, таких, как мониторинг строительных объектов с помощью БПЛА,

съемки панорам улиц из движущегося транспорта для выявления структурных дефектов, анализ архитектурных проектов для оценки соответствия проектным требованиям, и аэрофотосъемка для картографирования и мониторинга земных изменений [6]. Например, при анализе видеопотока с дронов в условиях аэрофотосъемки важно своевременно обнаружить и устранить смазанные кадры для точного создания ортофотопланов и трехмерных моделей местности [3]. Разработанные методы, такие, как улучшение обнаружения лиц в размытой видеозаписи с использованием суперразрешения, позволят повысить точность и надёжность систем видеонаблюдения. Автоматическое обнаружение смазанных кадров необходимо для улучшения качества обработки и анализа данных.

Таким образом, разработка эффективных методов для обнаружения и устранения смазанных кадров представляет собой важную задачу для обеспечения высококачественного анализа и обработки видеоинформации.

В данной статье рассматриваются различные подходы к обнаружению смазанных кадров. Авторами предлагается оригинальное решение, использующее нейросетевые методы для повышения точности определения смазанных кадров. В частности, предлагается применять сверточные нейронные сети для анализа спектра изображений, представленного в пространстве Хафа. Такой подход сочетает в себе преимущества спектрального анализа и преобразования Хафа [7], что позволяет более эффективно извлекать признаки, характерные для смазанных кадров.

Описание алгоритма

На первом этапе мы получаем спектральный анализ каждого кадра. Спектральный анализ используется для изучения частотных характеристик изображения [8]. Смазанные кадры, как правило, характеризуются пониженным содержанием высокочастотных компонент. Это происходит

потому, что смазывание в пространственной области приводит к сглаживанию деталей изображения и уменьшению резкости, что в частотной области соответствует уменьшению амплитуд высокочастотных составляющих.

Сравнивая спектры чистых и смазанных изображений на рис. 1, вызванных эффектом движения камеры или объекта, можно наблюдать характерные признаки смазывания в частотной области.

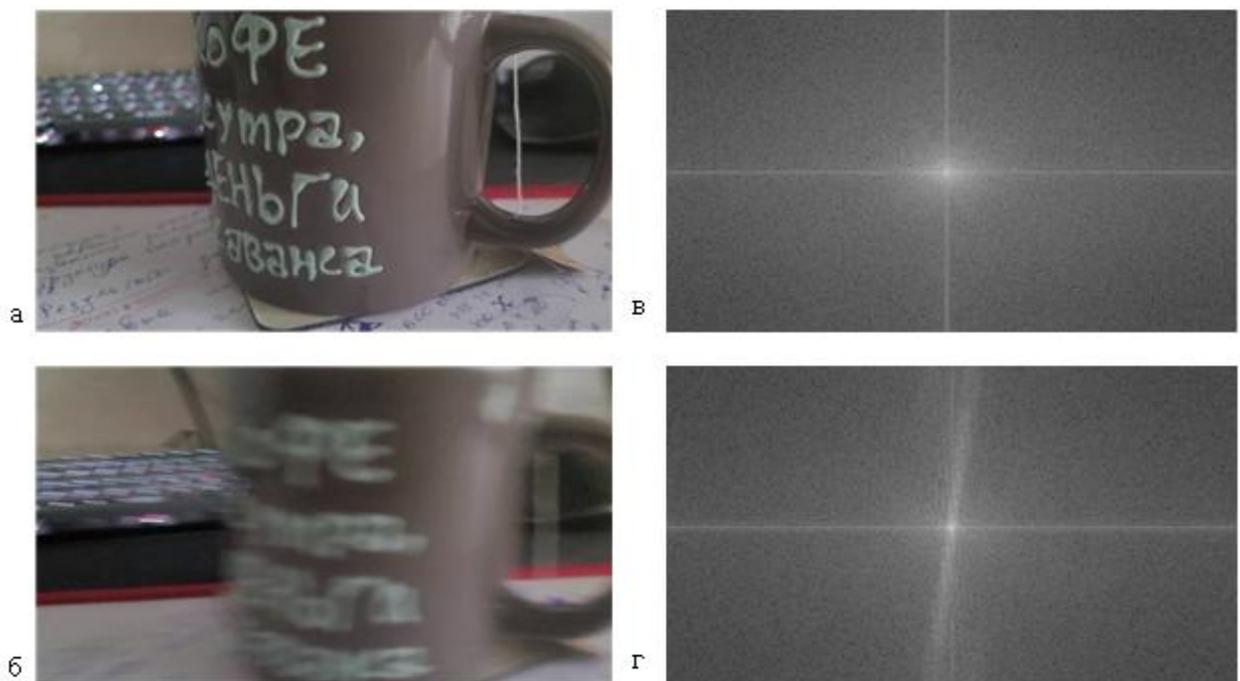


Рис. 1. – Изображения и их спектры (а – чистое изображение, б – смазанное изображение, в – спектр чистого изображения, г – спектр смазанного изображения)

Анализируя спектр размытого изображения, мы наблюдаем явно выделенную область, близкую к своей форме к прямой, что является отличительным признаком размытого кадра в видеопотоке.

На втором этапе полученные спектры представляются в пространстве Хафа. Преобразование Хафа — это метод обработки изображений, используемый для выявления геометрических фигур, таких, как прямые

линии [7], окружности и эллипсы, на изображениях. Основная идея заключается в преобразовании точек изображения в параметрическое пространство, где легче выявлять данные фигуры.

В контексте обнаружения смазанных кадров пространство Хафа используется для выявления линий и границ, которые в смазанных кадрах часто нечеткие и трудно различимы. Неразмытое изображение будет иметь более высокие объемы и выраженные пики в преобразовании Хафа в определенных угловых диапазонах.

Преобразование Хафа для двух видов изображений представлено на рис. 2.

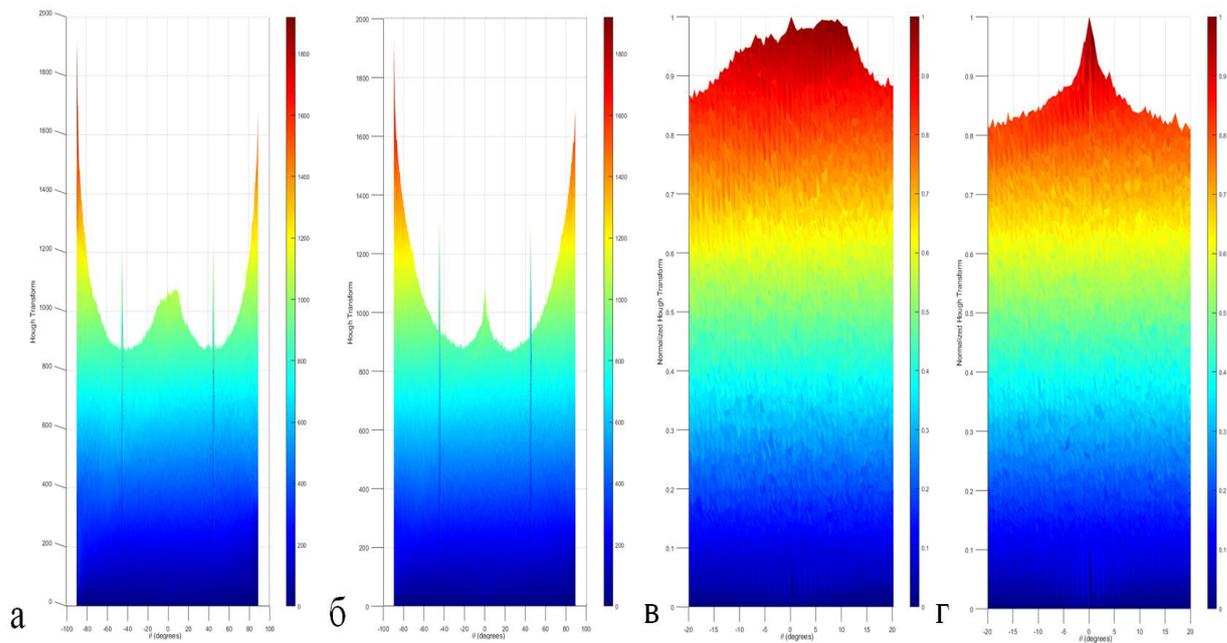


Рис. 2. – Преобразование Хафа (а – размытие движением камеры, б – чистое изображение, в – нормализованное преобразование Хафа «а» в диапазоне углов $\theta[-20:20]$, г – нормализованное преобразование Хафа «б» в диапазоне углов $\theta[-20:20]$)

Нормализованные объемы для различных угловых диапазонов в преобразовании Хафа показывают, что размытое изображение (рис. 2в) имеет больший объем в сравнении с чистым изображением.

Преобразование Хафа чистого изображение (рис. 2г) имеет более высокие пиковые значения в диапазоне углов θ $[-20:20]$, что свидетельствует о наличии четких и выраженных линий.

Для разработки автоматического детектора после представления изображения в пространстве Хафа было решено использовать нейросетевую модель для детектирования смазанных кадров. Нейросетевая модель базируется на сверточной нейронной сети с нижеследующими параметрами.

Таблица № 1

Параметры сверточной нейронной сети

Слой (тип)	Выходная форма слоя	Параметры
dense (Dense)	(None, 256)	530842624
dropout (Dropout)	(None, 256)	0
dense_1 (Dense)	(None, 128)	32896
dropout (Dropout)	(None, 128)	0
dense_2 (Dense)	(None, 3)	387
Всего параметров	530875907	
Обучаемые параметры	530875907	

Для обучения нейронной сети был использован датасет, состоящий из 5000 кадров, разделенных на тестовую и проверочную выборки в соотношении 30% и 70% соответственно. Датасет включал в себя как четкие, так и смазанные кадры.

Таким образом, финальный алгоритм можно представить следующим образом:

Этап 1. Предварительная обработка: Выделение яркостной компоненты из входного изображения.

Этап 2. Спектральный анализ и преобразование Хафа: кадры представляются в частотной области с использованием быстрого преобразования Фурье [9]. Из полученного спектра извлекаются

характеристики, связанные с высоким содержанием частот. Спектры кадров подвергаются преобразованию Хафа для выявления линий и границ. Выделенные линии анализируются на предмет их четкости и структуры.

Этап 3. Классификация: извлеченные признаки поступают на вход нейросетевой модели, которая принимает решение о смазанности кадра.

Перед обучением данные нормализуются с помощью StandardScaler. Обученная модель сохраняется для дальнейшего использования. Нейронная сеть состоит из трех полносвязных слоев с активацией ReLU и Dropout для регуляризации. Выходной слой использует активацию softmax для многоклассовой классификации.

Размер нейросетевой модели составляет 6 ГБ. Модель была протестирована на персональном компьютере с операционной системой Windows 10, оснащённом видеокартой Nvidia RTX 3070 Ti, а также на микрокомпьютере Nvidia Jetson Xavier NX, оснащённом 384-ядерным видеоядром NVIDIA Volta™ GPU с 48 тензорными ядрами [10]. Точность обученной нейронной сети составляет 92%.

Оценка достоверности предложенного решения

Для дополнительной оценки эффективности предложенного авторского решения было проведено сравнение существующих методов и алгоритмов, которые могут быть использованы при решении поставленной задачи, а именно - метода Лапласиана и метода ручной выборки.

Оператор Лапласа, являющийся оператором второго порядка, усиливает изменения интенсивности пикселей и выявляет края на изображениях. Смазанные кадры характеризуются менее выраженными краями, что приводит к снижению значений Лапласиана. В основе алгоритма выделения смазанных кадров с использованием оператора Лапласиана лежит анализ гистограммных значений пикселей.

Недостаток использования оператора Лапласа для обнаружения смазанных кадров заключается в его чувствительности к шуму. Поскольку Лапласиан усиливает все изменения интенсивности пикселей, он может также усилить шум на изображении, что может привести к ложным срабатываниям и снижению точности обнаружения смазанности. Это требует дополнительных шагов по предварительной обработке изображений, таких, как сглаживание или фильтрация, для уменьшения воздействия шума.

Алгоритм, основанный на операторе Лапласа, оказался эффективным только в 65% случаев.

Ручной метод включает в себя привлечение оператора для поиска и идентификации смазанных кадров. Основная особенность этого метода заключается в его высокой трудоемкости и низкой производительности, что ограничивает его эффективность в долгосрочной перспективе.

Ручной метод оказался самым эффективным (около 98%), но и самым продолжительным по времени. Также стоит отметить, что выбор искаженных кадров полностью зависит от оператора и не всегда бывает корректным (ошибка около 2%).

Заключение

В данной статье был представлен метод обнаружения смазанных кадров, основанный на использовании нейросетевой модели, анализирующей изображения в частотной области в пространстве Хафа. Проведенное исследование показало, что предложенный метод обладает высокой эффективностью и точностью в выявлении смазанных кадров, что делает его перспективным для применения в различных областях. Особенно актуально его использование в системах видеонаблюдения, медицинской визуализации и аэрофотосъемки, где четкость и качество изображений критически важны для принятия правильных решений. Полученные результаты подтверждают возможность интеграции разработанного метода в существующие системы

анализа изображений, что позволит значительно повысить надежность и точность обработки визуальной информации в условиях, требующих высокой степени детализации.

В будущем планируется расширение набора данных и улучшение архитектуры модели для достижения еще более высоких результатов.

Литература

1. Singh, N., Rathore, S.S. & Kumar, S. Towards a super-resolution based approach for improved face recognition in low-resolution environment. *Multimed Tools Appl* 81, 38887–38919. 2022. URL: doi.org/10.1007/s11042-022-13160-z.

2. Khosravi, Mohamadreza & Tavallali, Pooya. (2021). Real-time statistical image and video processing for remote sensing and surveillance applications. *Journal of Real-Time Image Processing*, 18. URL: doi.org/10.1007/s11554-021-01168-x.

3. Menaka, K., Yogameena, B., Nagananthini, C. 2017. Improving Face Detection in Blurred Videos for Surveillance Applications. In: Raman, B., Kumar, S., Roy, P., Sen, D. (eds) *Proceedings of International Conference on Computer Vision and Image Processing. Advances in Intelligent Systems and Computing*, vol 460. Springer, Singapore. URL: doi.org/10.1007/978-981-10-2107-7_14.

4. Grega, M., Leszczuk, M., Duplaga, M., Fraczek, R. (2010). Algorithms for Automatic Recognition of Non-informative Frames in Video Recordings of Bronchoscopic Procedures. In: Piętka, E., Kawa, J. (eds). *Information Technologies in Biomedicine. Advances in Intelligent and Soft Computing*, vol 69. Springer, Berlin, Heidelberg. URL: doi.org/10.1007/978-3-642-13105-9_53.

5. Саенко, А., Габов, В. Устранение смазанных кадров при решении задач фотограмметрии // *Актуальные аспекты научных исследований*. М.: Международный научно-издательский центр «Твоя наука», 2024. С. 52-59.

6. Пивень, П. В. Трансформации горно-лесного пояса Алтая вследствие климатических изменений // *Комплексное изучение экосистем горных*

территорий: Сборник материалов VI Кавказского Международного экологического форума, Грозный, 20–21 октября 2023 года. – Грозный: Чеченский государственный университет им. А.А. Кадырова, 2023. – С. 274-277. URL: doi.org/10.36684/102-1-2023-274-277.

7. Duda, R. O., & Hart, P. E. (1972). Use of the Hough Transformation to Detect Lines and Curves in Pictures. *Communications of the ACM*, 15(1), pp. 11-15.

8. Gonzalez, R. C., & Woods, R. E. (2002). *Digital Image Processing*. Prentice Hall. 793 p.

9. Oppenheim, A. V., & Schaffer, R. W. (2009). *Discrete-Time Signal Processing*. Prentice Hall. 1120 p.

10. NVIDIA Jetson Xavier. – URL: nvidia.com/en-us/autonomous-machines/embedded-systems/jetson-xavier-nx/ (accessed 12/05/24).

References

1. Singh, N., Rathore, S.S. & Kumar, S. Towards a super-resolution based approach for improved face recognition in low-resolution environment. *Multimed Tools Appl* 81, 38887–38919. 2022. URL: doi.org/10.1007/s11042-022-13160-z.

2. Khosravi, Mohamadreza & Tavallali, Pooya. 2021. Real-time statistical image and video processing for remote sensing and surveillance applications. *Journal of Real-Time Image Processing*, 18. URL: doi.org/10.1007/s11554-021-01168-x.

3. Menaka, K., Yogameena, B., Nagananthini, C. (2017). Improving Face Detection in Blurred Videos for Surveillance Applications. In: Raman, B., Kumar, S., Roy, P., Sen, D. (eds). *Proceedings of International Conference on Computer Vision and Image Processing*. *Advances in Intelligent Systems and Computing*, vol 460. Springer, Singapore. URL: doi.org/10.1007/978-981-10-2107-7_14.

4. Grega, M., Leszczuk, M., Duplaga, M., Fraczek, R. (2010). Algorithms for Automatic Recognition of Non-informative Frames in Video Recordings of

Bronchoscopic Procedures. In: Piętka, E., Kawa, J. (eds) Information Technologies in Biomedicine. Advances in Intelligent and Soft Computing, vol 69. Springer, Berlin, Heidelberg. URL: doi.org/10.1007/978-3-642-13105-9_53.

5. Saenko, A., Gabov, V. Ustranenie smazannykh kadrov pri reshenii zadach fotogrammetrii. Aktual'nye aspekty nauchnykh issledovaniy. Moskva: Mezhdunarodnyy nauchno-izdatel'skiy tsentr «Tvoya nauka», 2024. pp. 52-59.

6. Piven', P. V. Transformatsii gorno-lesnogo poyasa Altaya vsledstvie klimaticheskikh izmeneniy. Kompleksnoe izuchenie ekosistem gornyykh territoriy: Sbornik materialov VI Kavkazskogo Mezhdunarodnogo ekologicheskogo foruma. Groznyy: Chechenskiy gosudarstvennyy universitet im. A.A. Kadyrova, 2023. pp. 274-277. URL: doi.org/10.36684/102-1-2023-274-277.

7. Duda, R. O., & Hart, P. E. (1972). Use of the Hough Transformation to Detect Lines and Curves in Pictures. Communications of the ACM, 15(1), pp. 11-15.

8. Gonzalez, R. C., & Woods, R. E. 2002. Digital Image Processing. Prentice Hall. 793 p.

9. Oppenheim, A. V., & Schafer, R. W. 2009. Discrete-Time Signal Processing. Prentice Hall. 1120 p.

10. NVIDIA Jetson Xavier. URL: nvidia.com/en-us/autonomous-machines/embedded-systems/jetson-xavier-nx/ (accessed 12/05/24).

Дата поступления: 1.08.2024

Дата публикации: 15.09.2024