Приложение 2

|  |
| --- |
| к приказу  Южно-Уральского государственного  университета |

от \_\_\_\_\_\_\_\_\_\_ г.\_ №\_\_\_\_\_\_\_\_\_\_\_\_\_\_

#### ТРЕБОВАНИЯ К ОФОРМЛЕНИЮ СТАТЬИ

Рукопись объемом не более 8 полных страницы ф. А-4 (включая УДК (ББК), аннотацию, ключевые слова, рисунки, графики, таблицы, библиографический список и т.д.) представляется отредактированная, в виде распечатанного текста с обязательным представлением электронного носителя с текстом. Правка текстов от руки не допускается.

Текст должен быть набран в редакторе MicrosoftWord на листах стандартного формата А4 (21,0×29,7 см). Параметры страницы: верхнее поле – 2,5 см, нижнее поле – 2,8 см (расстояние от края листа до нижнего колонтитула – 2,0 см), левое поле – 2,5 см, правое поле – 2,5 см. Во избежание трудностей последующего форматирования параметры страниц рукописи необходимо задавать до начала набора текста.

УДК (ББК) проставляется по левому краю (шрифт Times New Roman 14 пт). Название статьи печатается прописными буквами по центру строки; точку в конце заголовка не ставить (шрифт TimesNewRoman 14 пт, жирный). Фамилии авторов набираются строчными буквами (малыми) под названием статьи справа, не указывая при этом ученой степени и ученого звания автора, инициалы размещать перед фамилией; если авторов несколько, фамилии печатать в одну строку через запятые (шрифт TimesNewRoman 14 пт, курсивный). Основной текст набирается шрифтом TimesNewRoman 14 пт с одинарным межстрочным интервалом, красная строка – 0,7 см.

Аннотация и ключевые слова набираются шрифтом TimesNewRoman 13 пт (смещение относительно левого поля – 3 см, красная строка – 0,7 см).

Между названием статьи, авторами, аннотацией и текстом статьи пропустить по одной пустой строке. Необходимо задать автоматический перенос слов.

Недопустимы в оригиналах висячие строки, т. е. неполные строки в начале страницы. При подготовке статьи *не следует* использовать разрывы строк, более одного пробела подряд, явную расстановку переносов с помощью дефисов и т. п.

В формулах размер и начертание математических символов должны совпадать с размером и начертанием этих же символов внутри основного текста.

Подрисуночные подписи и таблицы набираются шрифтом TimesNewRoman 13 пт. На все рисунки и таблицы должны быть ссылки в тексте статьи. Допускаются цветные графики, схемы, диаграммы, прочий иллюстративный материал.

При оформлении рукописи необходимо руководствоваться «Инструкцией о порядке подготовки и издании внутривузовской литературы» (приказ № 40 от 25.02.2013) (п. 8.9 «Некоторые особенности набора текстового материала», с. 27–29).

Библиографический список составляется в соответствии с ГОСТ 7.1–2003. На все литературные источники, приведенные в библиографическом списке, в тексте статьи должны быть сделаны ссылки в порядке их упоминания.

Номера страниц не проставляются.

*Авторы обязаны подписать представляемый экземпляр рукописи.*

УДК 004.896

**АЛГОРИТМЫ ШУМОПОДАВЛЕНИЯ НА ВИДЕОРЯДЕ   
В РЕАЛЬНОМ ВРЕМЕНИ**

***П.П. Петров***

В настоящее время шумоподавление широко используется в алгоритмах предварительной обработки. Изображение ввиду технического устройства камеры и внешних помех может иметь зашумление, мешающее беспрепятственному наблюдению за целями.

Предполагается использовать, идеи архитектур нейронных сетей, таких, как FastDVDnet, BasicVSR++, ViDeNN.

В статье показывается, что глубокое обучение, в частности сверточные нейронные сети, способны извлекать важные характеристики из необработанного видеоряда в реальном времени Составлены сравнительные таблицы рассматриваемых методов, а также с информацией о наборах данных, использованных в рассмотренных статьях.

Ключевые слова: шумоподавление на видеоряде, компьютерное зрение, реальное время, нейронные сети.

**Актуальность.** Изображение ввиду технического устройства камеры и внешних помех может иметь зашумление, мешающее беспрепятственному наблюдению за целями. Видеокамеры сегодня применяются во множестве областях.

Одним из направлений развития является шумоподавление в реальном времени на видеоряде. Для решения этих задач используют методы машинного обучения и нейронных сетей. Их область применения охватывает многие области ТЗ: растениеводство, животноводство, оптимизация парка сельхозмашин и др. Их использование требует адаптации к природным и агроклиматическим условиям регионов, в которых расположены хозяйства. Для этого следует подготовить тренировочный набор данных для обучения сети. Глубокое обучение, а, в частности, сверточные нейронные сети, способны извлекать основные признаки из кадров.

**1. Сеть BasicVSR++ – основная идея состоит в том, чтобы использовать оптический поток в качестве базовых смещений** (рис. 1). Рассмотрим работу, в которой изучен способ использования оптического потока в качестве базовых смещений [1].

Поскольку использование временной информации при восстановлении видео является общей целью, предполагается, что успех базового VSR++ не ограничивается сверхразрешением видео. В этой работе расширяем сферу нашей деятельности на большее количество задач по восстановлению видео и представляем универсальный фреймворк, построенный на базовом VSR++. Для задач, где входные данные и выходные данные имеют одинаковое разрешение, уменьшается входное разрешение с помощью ступенчатых сверток для поддержания эффективности. В дополнение к сверхразрешению видео и улучшению качества сжатого видео, показано, что базовый VSR++ применим для устранения размытия и шумоподавления видео, достижение многообещающих показателей при высокой эффективности.

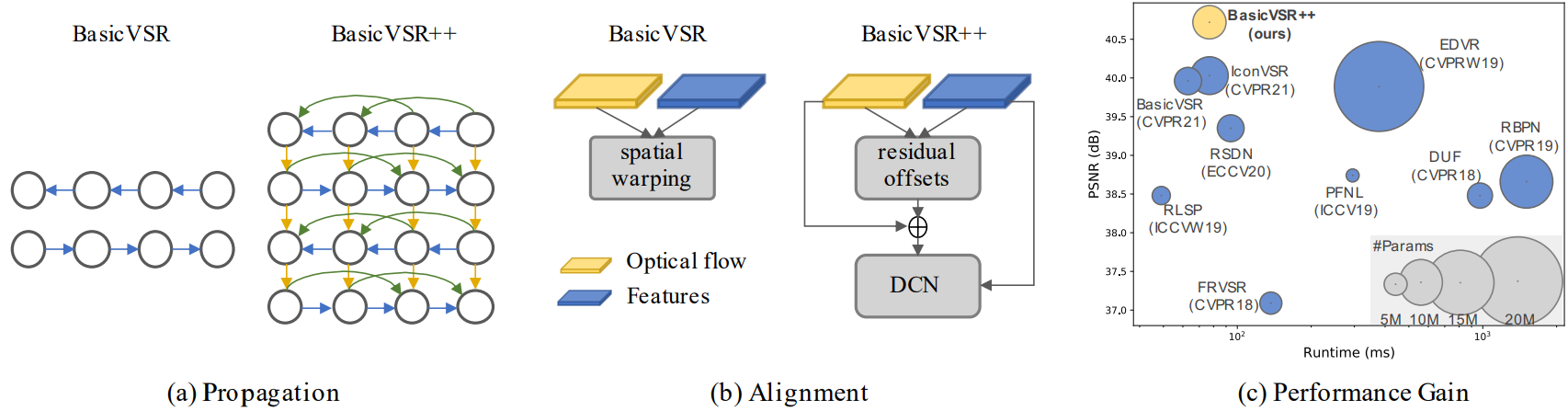


Рис. 1. Архитектура BasicVSR++ в сравнении с BasicVSR

Таким образом, данная работа представляет собой хорошую базу для проведения новых экспериментов, так как управляемое потоком деформируемое выравнивание в базовом VSR++ обеспечивает средство для более надежного выравнивания объектов в несоосных кадрах.

**2. FastDvdNet: использование последовательных кадров для подавления шумов в среднем кадре** (рис. 2, 3). Использование объемных (то есть пространственно-временных) соседей означает, что при подавлении шумов для данного пикселя (или патча) алгоритм будет искать похожие пиксели

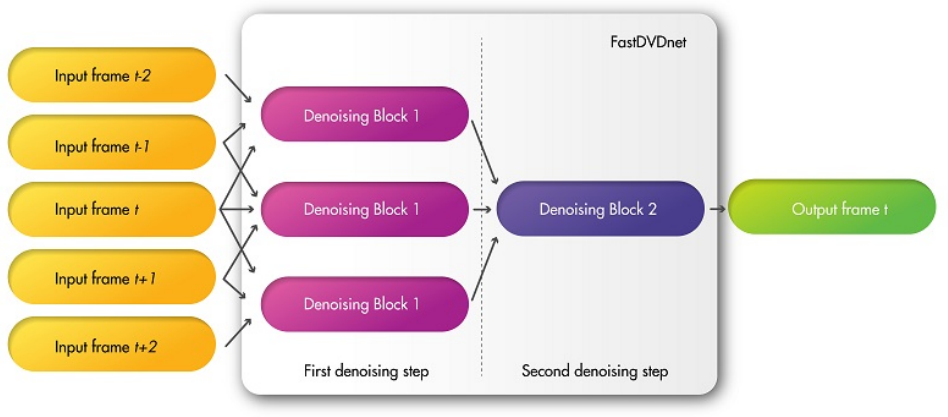


Рис. 2. Архитектура FastDVDNet

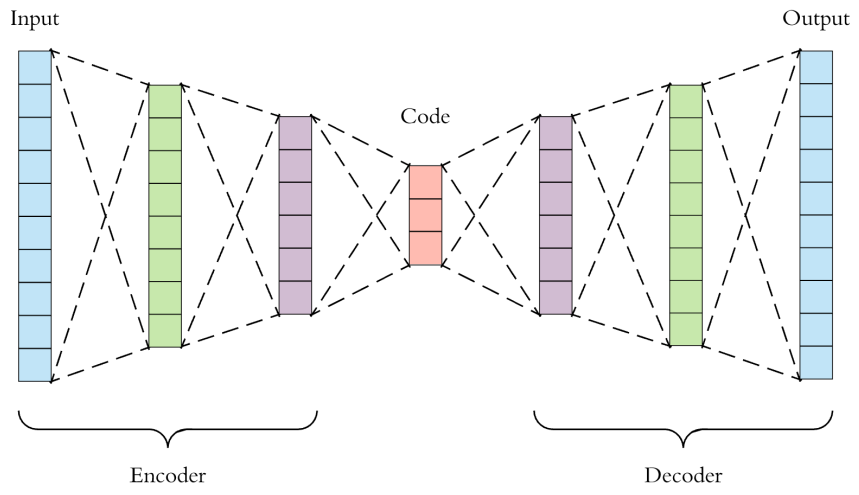


Рис. 3. Архитектура Denoising Block

(патчи) не только в эталонном кадре, но также в соседних кадрах последовательности. Преимущества этого подхода двояки. Во-первых, временные соседи предоставляют дополнительную информацию, которую можно использовать для подавления шумов на эталонном кадре. Во-вторых, использование временных соседей помогает уменьшить мерцание, так как остаточная ошибка в каждом кадре будет коррелирована [2, 3].

**3. ViDeNN – объединение двух сетей, выполняющих сначала пространственное шумоподавление одного кадра, а затем временное шумоподавление в окне из трех кадров** (рис. 4). Каждый кадр будет проходить через пространственный шумоподавитель CNN. Временный CNN принимает в качестве входных данных три пространственно сглаженных кадра и выводит окончательную оценку центрального кадра. Оба канала CNN сначала оценивают остаточный шум, т. е. нежелательные значения, которые шум добавляет к изображению, а затем вычитают их из зашумленного входного сигнала (⊕ означает сложение двух сигналов, а «–» отрицание). ViDeNN состоит только из сверточных слоев. Количество карт объектов указано в нижней части каждого слоя [4].

Эффективность шумоподавления CNN с пространственным шумоподавлением в значительной степени зависит от обучающих данных. Реальное распределение шума отличается от гауссовского, поскольку оно не является чисто аддитивным [5], но содержит часть, зависящую от сигнала. По этой причине модели CNN обучались только аддитивному белому гауссову шуму (AWGN) не удается сгладить шумы изображений реального мира [6]. Наша цель – достичь хорошего баланса между производительностью и гибкостью, обучая единую сеть нескольким моделям шума. Как показано в табл. 1, наш Spatial-CNN может обрабатывать слепые Гауссовское шумоподавление: продолжается исследование его возможности обобщения, представив модель шума, зависящую от сигнала. Эта конкретная модель шума в уравнении 1 состоит из двух основных составляющих: шума фотонного выстрела (PSN) и шум при считывании. PSN является основным источником шума в условиях низкой освещенности, где Nsat учитывает количество насыщенных электронов [7].

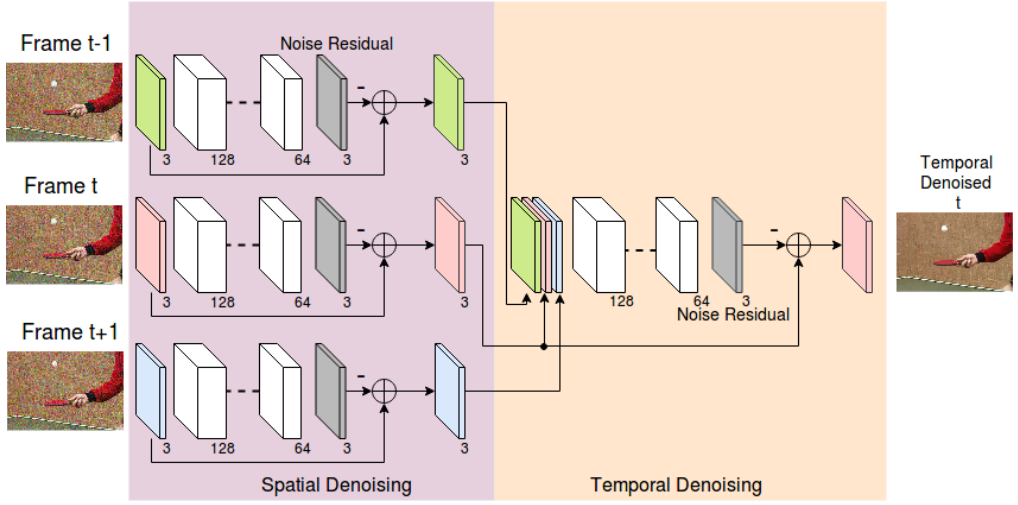


Рис. 4. Архитектура ViDeNN

**4. Сеть RTD** (рис. 5). Это новый метод денойзинга на основе глубокого обучения для реального времени с возможностью управления. RTD включает в себя три основных компонента:

1. Основная сеть, т. е. : → , генерирует несколько карт шума с фиксированным уровнем, где L – количество предопределенных уровней шума [6].
2. Блок Noise Decorrelation (ND), который обеспечивает возможность редактирования сгенерированных карт шума [8].
3. Модуль AutoTune, который предоставляет набор параметров управления по умолчанию для получения лучшего результата денойзинга.

Из особенностей данной работы стоит отметить скорость работы модели, была достигнута поддержка работы в 20+ Гц. Кроме RGB на вход модель получает еще 13 карт признаков, вычисляемых из RGB изображения [9]. Подчеркивается, что такой подход позволяет достигнуть большей обобщающей способности при использовании маленького набора данных.

**Сравнение моделей.** В табл. 1 приведены точность работы рассмотренных выше методов, а также используемые в них архитектуры, в соответствии с порядком их рассмотрения.

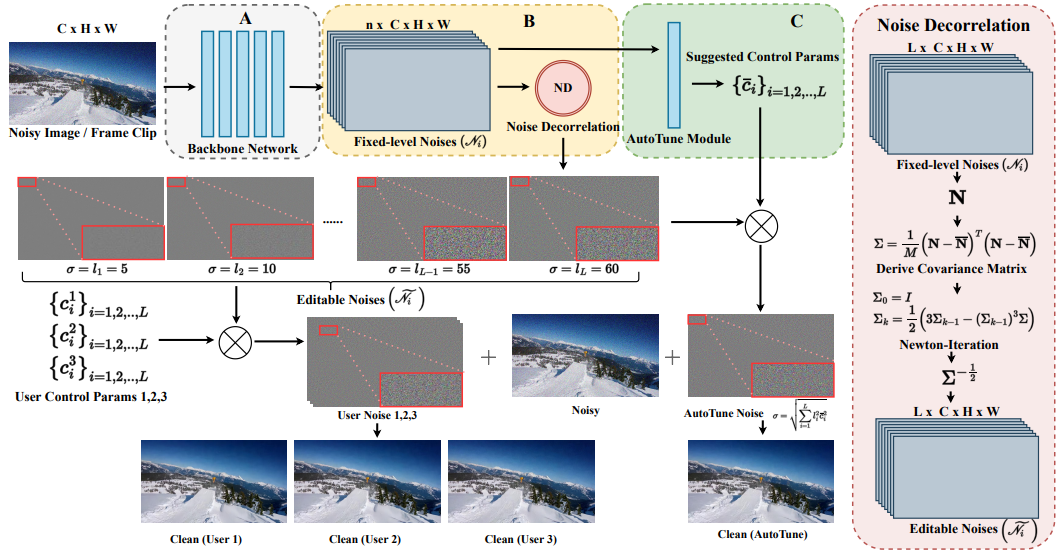


Рис. 5. Архитектура нейронной сети RTD

Таблица 1

Рассмотренные методы шумоподавления

|  |  |  |  |
| --- | --- | --- | --- |
| Название сети | Скорость  обработки кадра (с) | Шум (σ) | Конечный PSNR |
| Сеть BasicVSR++ | 0.24 | 50 | 34.17 |
| Сеть ViDeNN | 2.4 | 50 | 26.17 |
| Сеть FastDvdNet  Сеть RTD | 0.1  2.5 | 50  50 | 28.5 |

В первой строке табл. 1 отображено, что лучшей сетью, для данного эксперимента, оказалась FastDVDNet. Сравнение метрик точности для моделей не будет являться объективным, так как они обучались и тестировались на разных наборах данных. Данное сравнение является темой будущих исследований [10].

**Используемые наборы данных.** В FastDVDNet разработчики не описывают набор данных для обучения их сети.

**Заключение.** Прямое сравнение качества работы моделей по метрикам, представленным в статьях, не будет являться объективным из-за различия   
в используемых наборах данных. На основании этого можно прийти к выводу о необходимости переобучить либо дообучить модели на одном наборе данных для их прямого сравнения.

На основании табл. 1 делаем вывод, что сеть FastDVD является одной из лучших в области шумоподавления. Также существует актуальность поиска решений для улучшения качества работы FastDVD.

Библиографический список

1. On the Generalization of BasicVSR++ to Video Deblurring and Denoising / K.C.K. Chan, S. Zhou, X. Xu, C.C. Loy. – arXiv:2204.05308v2 [cs.CV] 19 Jun 2022.

2. Tassano, M. FastDVDnet: Towards Real-Time Deep Video Denoising Without Flow Estimation / M. Tassano, J. Delon, T. Veit. – arXiv:1907.01361v2 [cs.CV] 29 Apr 2020.

3. LeCun, Y.A. Convolutional networks for images, speech, and time series /   
Y.A. LeCun, Y. Bengio. – 1998.

4. Claus, M. ViDeNN: Deep Blind Video Denoising / M. Claus, J. Gemert. –   
arXiv:1904.10898v1 [cs.CV] 24 Apr 2019.

5. Real-time Controllable Denoising for Image and Video / Zhaoyang Zhang,   
Yitong Jiang, Wenqi Shao et al. This CVPR paper is the Open Access version, provided by the Computer Vision Foundation.

6. Deep residual learning for image recognition / He, K., Zhang, X., Ren, S. et al. // Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2016-Decem. – 2016. – P. 770–778. DOI: 10.1109/ CVPR.2016.90

7. Погонышев, В.А. Нейронные сети в цифровом сельском хозяйстве / В.А. Погонышев, Д.А. Погонышева, В.Е. Ториков // [Вестник Брянской ГСХА](https://elibrary.ru/contents.asp?id=46696474). – 2021. – № [5 (87)](https://elibrary.ru/contents.asp?id=46696474&selid=46696486). – С. 68–71.

8. U-Net: Convolutional Networks for Biomedical Image Segmentation / O. Ron-neberger, P. Fischer, T. Brox // Navab N., Hornegger J., Wells W.M. (Eds.) Medical   
Image Computing and Computer-Assisted Intervention – MICCAI 2015. – Springer International Publishing, Cham, 2015. – Vol. 9351. – P. 234–241.

9. Deep learning-based early weed segmentation using motion blurred UAV images of sorghum fields / N. Genze, R. Ajekwe, Z. Güreli et al. // Computers and Electronics in Agriculture. – November 2022. – Vol. 202, 107388. DOI: 10.17632/4hh45vkp38.4

10. Rethinking Atrous Convolution for Semantic Image Segmentation / L.-C. Chen, G. Papandreou, F. Schroff et al. – 2017. – P. 2331–8422.