

**РАЗРАБОТКА НЕЙРОСЕТЕВОГО МЕТОДА
ШУМОПОДАВЛЕНИЯ ДЛЯ ВИДЕО В ФОРМАТЕ RAW
НА ОСНОВЕ РЕАЛИСТИЧНОГО НАБОРА ДАННЫХ,
СОБРАННОГО В УСЛОВИЯХ НИЗКОЙ
ОСВЕЩЕННОСТИ**

Железов Максим Сергеевич

Студент

Факультет ВМК МГУ имени М. В. Ломоносова, Москва, Россия

E-mail: maksim.zhelezov@graphics.cs.msu.ru

Научный руководитель — Ватолин Дмитрий Сергеевич

Задача устранения шума в RAW видео направлена на восстановление последовательности кадров, искажённых сенсорным шумом, особенно в условиях низкой освещённости. Входными данными алгоритма являются зашумлённые RAW видеопоследовательности, представляющие собой линейные измерения сигнала сенсора. В отличие от стандартных sRGB-данных, RAW формат сохраняет полный динамический диапазон сцены и информацию о статистике шума, что делает его более информативным, но одновременно более сложным для обработки. Выходными данными является восстановленная видеопоследовательность с подавленным шумом и сохранёнными пространственно-временными деталями.

В последние годы задача видеושумоподавления активно развивается, при этом всё большее внимание уделяется обработке данных непосредственно в RAW-пространстве. Это обусловлено тем, что линейный отклик сенсора позволяет корректно моделировать физические процессы формирования изображения и учитывать реальные шумовые характеристики. Однако при низкой освещённости уровень шума существенно возрастает, а полезный сигнал может быть сопоставим с уровнем шума или даже ниже его, что значительно усложняет задачу восстановления.

В данной работе предлагается метод устранения шума в RAW видеопоследовательностях, ориентированный на обработку данных непосредственно в линейном пространстве сенсора и учитывающий пространственно-временную корреляцию кадров. Входными данными метода является последовательность зашумлённых RAW кадров, представленных в виде мозаики Байера и нормализованных в соответствии с параметрами сенсора. Выходом является восстановленная видеопоследовательность с подавленным шумом и сохранённой

текстурой, мелкими деталями и временной согласованностью. Предлагаемый подход сочетает многоуровневую пространственную обработку и явное моделирование временных связей между соседними кадрами.

Предлагается пространственно-временная архитектура для подавления шума в RAW видеопоследовательностях, работающая непосредственно в линейном пространстве сенсора. На вход подаётся последовательность соседних RAW кадров (в формате RGGB), из которой восстанавливается центральный кадр. Каждый кадр проходит через модуль извлечения признаков на основе сверточных residual-блоков без нормализации, после чего выполняется обучаемое выравнивание соседних кадров с использованием deformable convolution [1]. Далее признаки агрегируются в пространственно-временном блоке, сочетающем 3D-свертки и механизм временного внимания, что позволяет эффективно учитывать межкадровые зависимости и подавлять области с нестабильным сопоставлением.

Объединённое представление обрабатывается многоуровневым энкодер–декодером U-Net [2] типа с расширенным рецептивным полем в bottleneck-слое, а восстановление выполняется по остаточному принципу — сеть предсказывает карту шума, которая вычитается из центрального кадра.

Экспериментальная оценка на наборе RAW видеоданных показала устойчивое улучшение объективных метрик качества (SSIM [3]) по сравнению с пространственными и базовыми временными методами, а также снижение временного мерцания в динамических сценах. Визуальный анализ подтверждает лучшее сохранение текстур и мелких деталей при сильном уровне шума.

Литература

1. Dai J., Qi H., Xiong Y., Li Y., Zhang G., Hu H., Wei Y. Deformable convolutional networks // In Proceedings of the IEEE international conference on computer vision, 2017, P. 764–773.
2. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation // In Medical Image Computing and Computer-Assisted Intervention, 2015, P. 234–241
3. Wang Z., Bovik A.C., Sheikh H.R., Simoncelli E.P. Image quality assessment: from error visibility to structural similarity // In IEEE transactions on image processing, 2004, P. 600–612.