

# JointLoss: описание и принцип работы

## Введение

Передача стиля — задача компьютерного зрения, направленная на создание новых объектов визуального искусства. Его цель — синтезировать образ, сочетающий в себе узнаваемые стилевые образцы изображения стиля и сохраняющий тему изображения контента.

Метод Vector-NST основан на векторной графике и работе с изображениями формата SVG (scalable vector graphics) [1, 2]. В процессе работы модели получались изображения с неровными и неаккуратными контурами, требующие сглаживания.

Требовалось сгладить контуры для лучшего визуального качества выходного изображения. Ранее для данной задачи использовался xing-loss, однако его применение не было достаточно эффективным [3]. Поэтому необходимо было разработать новую функцию потерь, позволяющую более эффективно сглаживать контуры векторных изображений.

## Описание работы

SVG-формат изображений представляет собой текстовый файл, описывающий изображения в виде геометрических примитивов: линий, кривых, фигур, текста и так далее [2]. Основным способом описания таких примитивов является кубическая кривая Безье, задающаяся четырьмя опорными точками [4]. Идеей для создания новой функции потерь послужило предположение, что основной причиной неровностей контуров изображения является особая комбинация опорных точек кривой, а именно случай образования этой кривой так называемого стыка — точки, в которой касательная к кривой не определена (рис. 1)

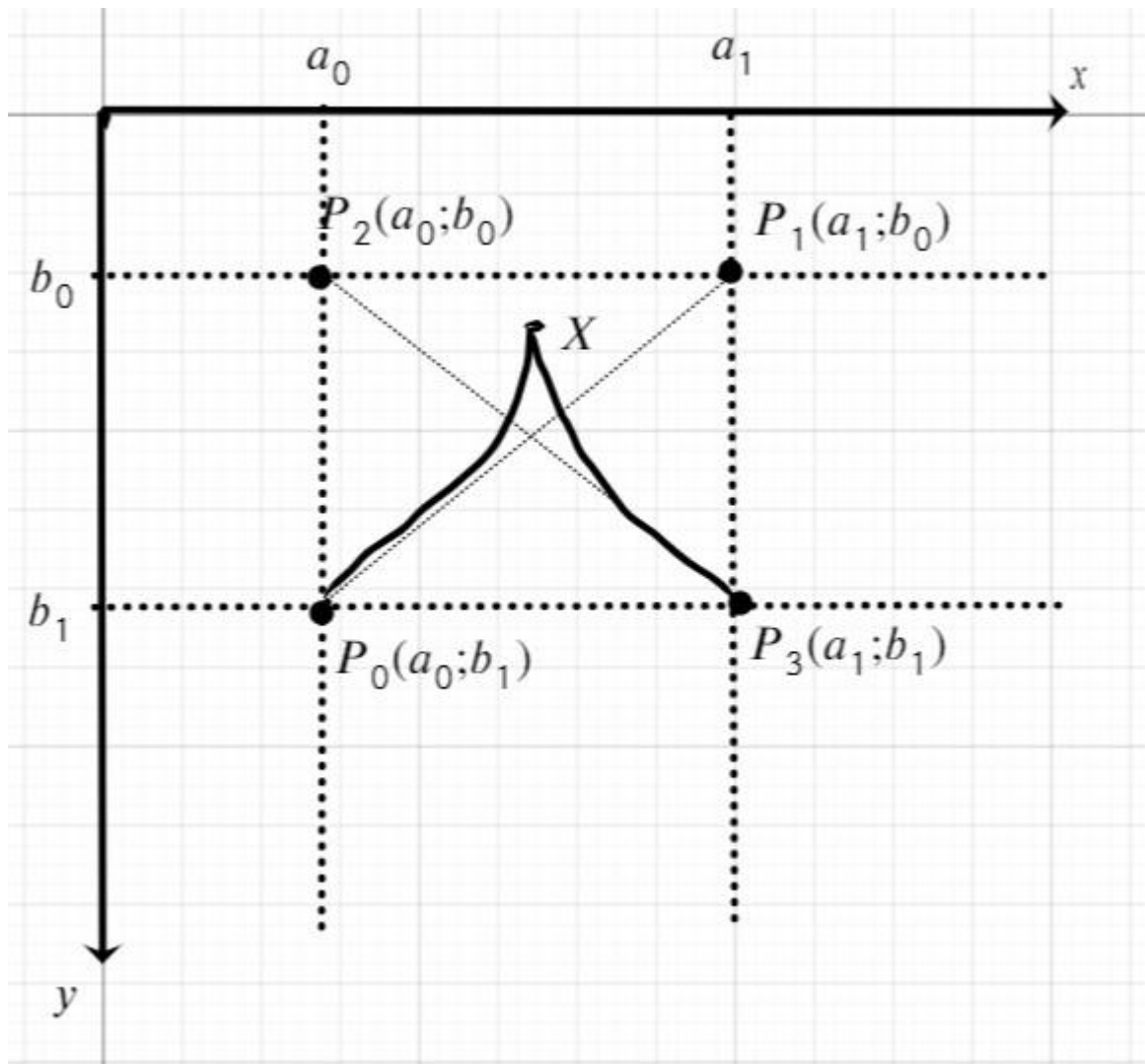


Рисунок 1 - Наличие стыка у кривой Безье

Поскольку кубическая кривая Безье является вектор-функцией в двумерном пространстве, то наличие такой особой точки вытекает из отсутствия производной в данной точке.

Пусть  $B(t) = (x(t), y(t)), t \in [0; 1]$  кривая Безье. Так как кривая кубическая, то она задаётся 4 точками:  $P_0(a_0; b_0), P_1(a_1, b_1), P_2(a_2, b_2)$  и  $P_3(a_3, b_3)$ . Тогда кривую можно описать функцией  $y(x)$  заданной параметрически:

$$x(t) = a_0 \cdot (1 - t)^3 + a_1 \cdot 3t(1 - t)^2 + a_2 \cdot 3t^2(1 - t) + a_3 \cdot t^3$$

$$y(t) = b_0 \cdot (1 - t)^3 + b_1 \cdot 3t(1 - t)^2 + b_2 \cdot 3t^2(1 - t) + b_3 \cdot t^3$$

При смене знаков и раскрытии скобок в некоторых местах получается следующая система:

$$x(t) = a_3 \cdot t^3 - 3a_2 \cdot (t^3 - t^2) + 3a_1 \cdot (t^3 - 2t^2 + t) - a_0 \cdot (t - 1)^3$$

$$y(t) = b_3 \cdot t^3 - 3b_2 \cdot (t^3 - t^2) + 3b_1 \cdot (t^3 - 2t^2 + t) - b_0 \cdot (t - 1)^3$$

Не всегда кривая может описываться именно функцией  $y(x)$ , так как при других координатах исходных точек получается вообще говоря не функция. Однако на данный момент стык наблюдался именно при представлении кривой в виде функции. Стоит отметить, что иногда функция, описывающая кривую, будет иметь вид не  $y(x)$ , а  $x(y)$ . Такое, например, возможно получить при замене местами точек  $P_0$  и  $P_1$  кривой на рис. 1. Данный момент будет объяснён подробнее ниже.

Итак, имея функцию  $y(x)$  заданную параметрически, следует найти точку, в которой касательная к кривой (графику функции) не существует. Рассмотрим производную данной функции:  $y'_x = \frac{y'_t}{x'_t}$ . Получим:

$$x'_t = 3a_3 \cdot t^2 - 3a_2 \cdot (3t^2 - 2t) + 3a_1 \cdot (3t^2 - 4t + 1) - 3a_0 \cdot (t^2 - 2t + 1)$$

$$y'_t = 3b_3 \cdot t^2 - 3b_2 \cdot (3t^2 - 2t) + 3b_1 \cdot (3t^2 - 4t + 1) - 3b_0 \cdot (t^2 - 2t + 1)$$

При делении одного уравнения на другое, сокращении и группировке слагаемых получается следующее уравнение (1):

$$\begin{aligned} y'_x &= \frac{b_3 \cdot t^2 - 3b_2 \cdot t^2 + 2b_2 \cdot t + 3b_1 \cdot t^2 - 4b_1 \cdot t + b_1 - b_0 \cdot t^2 + 2b_0 \cdot t - b_0}{a_3 \cdot t^2 - 3a_2 \cdot t^2 + 2a_2 \cdot t + 3a_1 \cdot t^2 - 4a_1 \cdot t + a_1 - a_0 \cdot t^2 + 2a_0 \cdot t - a_0} \\ &= \frac{(b_3 - 3b_2 + 3b_1 - b_0) \cdot t^2 + (2b_2 - 4b_1 + 2b_0) \cdot t + b_1 - b_0}{(a_3 - 3a_2 + 3a_1 - a_0) \cdot t^2 + (2a_2 - 4a_1 + 2a_0) \cdot t + a_1 - a_0} \end{aligned}$$

Данная производная не существует, когда знаменатель такой дроби равен 0, а числитель отличен от 0 (случай деления 0 на 0 стоит отметить отдельно). Однако в знаменателе и числителе полиномы второй степени от одной переменной, тогда можно найти корни этого полинома, учтя их существование в поле вещественных чисел. При обозначении  $k := a_2 - 2a_1 + a_0$ ,  $a := a_3 - 3a_2 + 3a_1 - a_0$ ,  $c := a_1 - a_0$  и  $D_1 := k^2 - ac$  и при выполнении условия

$k^2 \geq ac$ , получим формулу  $t_{1,2} = \frac{-k \pm \sqrt{D_1}}{a}$  (2). Если неравенство не выполняется, то производная функции существует всегда и знаменатель не обратится в 0). Остаётся проверить каждое из значений  $t_{1,2}$  на принадлежность отрезку  $[0; 1]$ , и тогда становится возможным находить точки, в которых образуется стык. По построению кривой Безье точек стыка возможно не более чем одна, а значит достаточно проверить что хотя бы одно из значений принадлежит отрезку.

Нахождение значений по формуле (2) ещё не гарантирует наличие стыка у кривой из-за возможного явления перегиба – математически это стремление значения производной к бесконечности (рис. 2)

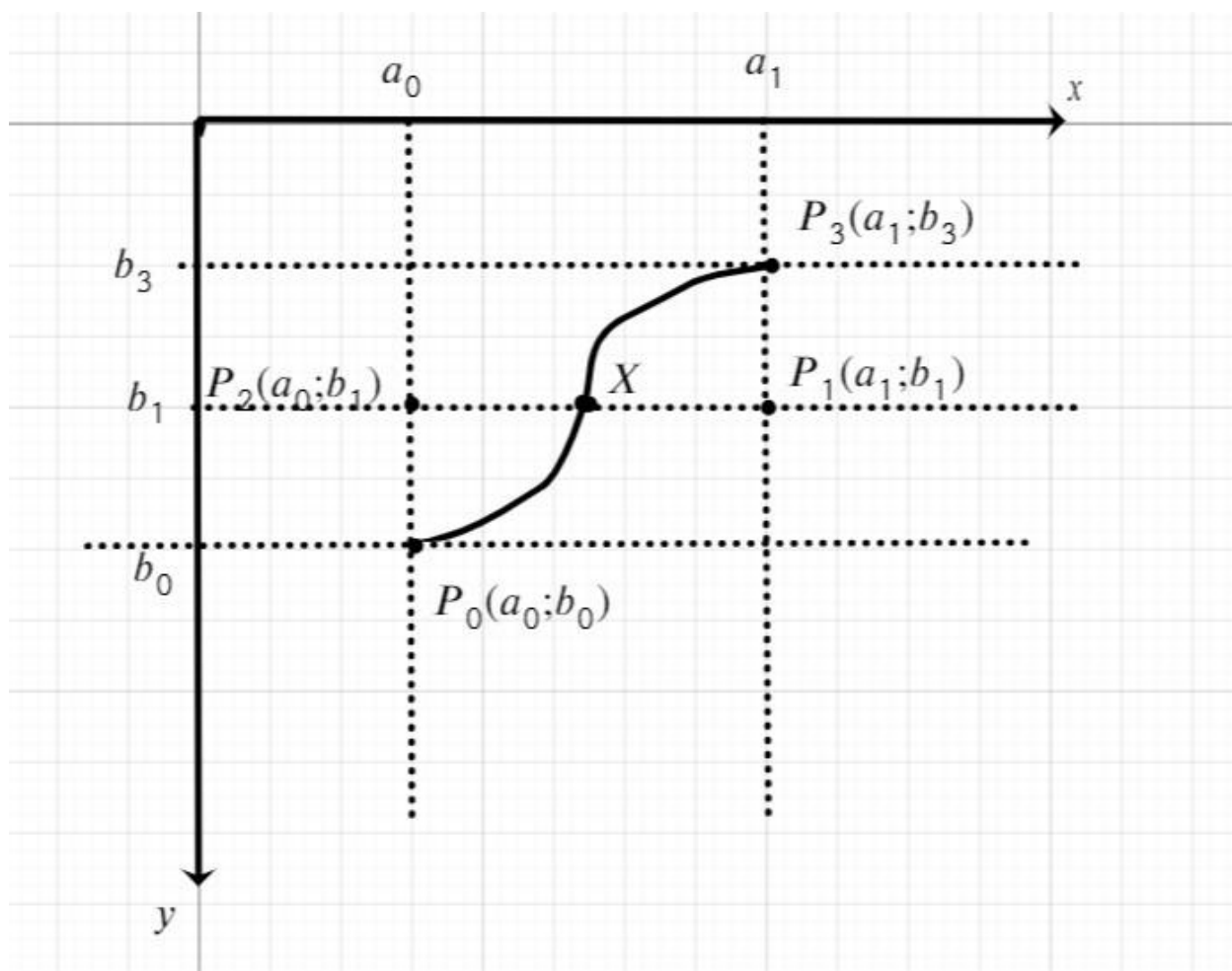


Рисунок 2 - наличие у кривой Безье перегиба

Здесь кривая Безье также описывается как функция  $y(x)$ , но у ней нет стыка, несмотря на несуществование производной в точке  $X$ . Однако, данного

явления можно избежать, если убедиться, что найденные по уравнению (2) значения соответствуют точкам локального экстремума функции, то есть, значения в точках  $t + \varepsilon, t - \varepsilon$  соответственно больше или меньше значения в точке  $t$ . Данное вычисление не составляет больших проблем, так как требуется лишь подставить нужные значения в качестве аргументов функции, описывающей кривую.

Теперь рассмотрим ситуацию, когда кривая задаётся не как функция  $y(x)$ , а как функция  $x(y)$ . В данном случае достаточно поменять координаты точек, выполнив как бы поворот координатной плоскости на  $90^\circ$ . В новом базисе можно рассматривать кривую уже как функцию  $y(x)$  и выполнять аналогичные действия, не забыв в конце вернуть координаты обратно в исходный базис. Таким образом, единственный случай, который не был рассмотрен, это одновременное равенство числителя и знаменателя дроби в уравнении (1) нулю. Следует отметить факт, что если при каком-то значении  $t = t_0$  и числитель и знаменатель обращается в 0, то данное значение является корнем каждого из полиномов и в таком случае числитель и знаменатель можно сократить на соответствующий множитель  $(t - t_0)$  и рассматривать другой корень знаменателя при его наличии.

Получив значение  $t$ , при котором у кубической кривой Безье возникает стык, можно рассчитать величину кривизны кривой, то есть насколько сильно данный стык влияет на неровность контура этой кривой. Но поскольку для самой точки стыка кривизна не имеет смысла, то следует посмотреть на кривизну кривой в  $\varepsilon$ -окрестности данной точки. Кривизна кривой в точке  $t$  высчитывается по формуле (3)  $k(t) = \frac{x'_t y''_t - x''_t y'_t}{(x'^2_t + y'^2_t)^{\frac{3}{2}}}$ , что позволяет без особых усилий использовать найденные ранее значения производных для расчёта функции потерь. Чем больше данное значение, тем более неровным получается контур кривой, что идеально подходит для сглаживания контуров с помощью нейронной сети.

## Заключение

Таким образом, полученные рассуждения предоставили возможность создать новую функцию потерь, которая нацелена на сглаживание контуров векторного изображения путём анализа кубических кривых Безье, входящих в его состав. С помощью представления кривой как вектор-функции можно определить наличие либо отсутствие у неё стыка, влекущего за собой неровность контуров. Дальнейшее улучшение работоспособности данной функции позволит значительно повысить качество изображений, получаемых в результате работы нейронной сети, что безусловно будет полезно при последующей обработке и анализе выходных данных.

## Список источников

- 1 – описание принципов работы vector neural style transfer модели:  
[https://www.researchgate.net/publication/369063548\\_Neural\\_Style\\_Transfer\\_for\\_Vector\\_Graphics](https://www.researchgate.net/publication/369063548_Neural_Style_Transfer_for_Vector_Graphics)
- 2 – документация svg-формата изображений <https://svgwg.org/svg-next/struct.html>
- 3 – описание работы использовавшейся ранее функции потерь  
[https://openaccess.thecvf.com/content/CVPR2022/papers/Ma\\_Towards\\_Layer-Wise\\_Image\\_Vectorization\\_CVPR\\_2022\\_paper.pdf](https://openaccess.thecvf.com/content/CVPR2022/papers/Ma_Towards_Layer-Wise_Image_Vectorization_CVPR_2022_paper.pdf)
- 4 – описание математической модели кривой Безье  
[https://ru.wikipedia.org/wiki/%D0%9A%D1%80%D0%B8%D0%B2%D0%B0%D1%8F\\_%D0%91%D0%B5%D0%B7%D1%8C%D0%B5](https://ru.wikipedia.org/wiki/%D0%9A%D1%80%D0%B8%D0%B2%D0%B0%D1%8F_%D0%91%D0%B5%D0%B7%D1%8C%D0%B5)