

# АНАЛИЗ ЭФФЕКТИВНОСТИ ФУНКЦИОНИРОВАНИЯ КОНВЕЙЕРНОЙ ВЫЧИСЛИТЕЛЬНОЙ СИСТЕМЫ ДЛЯ ПОИСКА ИЗОБРАЖЕНИЙ В БАЗЕ ДАННЫХ ПО ИХ СОДЕРЖИМОМУ

Костюкова Н.С.  
Кафедра ПМиИ ДонНТУ  
Shozda@r5.dgtu.donetsk.ua.

## **Abstract**

*Kostyukova N.S. Performance analysis for pipeline content-based image retrieval system. This article is devoted to the problem of performance analysis for specialized image retrieval system. Waiting time problem is described, pipeline optimization is performed.*

## **Введение**

Задача поиска изображений в больших базах данных по заданному образцу принадлежит к числу тех, которые не могут быть решены в реальном масштабе времени, несмотря на модификации используемых алгоритмов обработки. Одним из эффективных путей в решении указанной проблемы является использование вычислительной техники со специализированной архитектурой. Вопросы ускорения вычислений в результате применения архитектурных и аппаратных усовершенствований давно являются предметом исследований в различных областях, и, в частности, в сфере обработки изображений [1, 2]. Ранее автором анализировались возможности использования данного подхода в процессе контекстного поиска изображений как для решения задачи в целом, так и отдельных ее этапов [3, 4]. Целью данной работы является описание специализированной вычислительной системы для контекстного поиска изображений и решение задачи эффективной организации конвейера вычислений.

## **1. Специализированная вычислительной системы для поиска изображений**

При создании системы контекстного поиска изображений, работающей в режиме реального времени, очень важным является выбор рациональной (с точки зрения быстродействия и аппаратных затрат) структуры вычислительной системы в целом и ее составляющих. Ранее автором было выполнено исследование возможных вариантов построения как системы в целом, так и отдельных подсистем [3, 4]. При использовании

модифицированного алгоритма поиска, описанного в [5, 6], процесс контекстного поиска изображений состоит из следующих последовательно выполняющихся шагов: приведение цветов изображения-образца к базовому набору; построение ЦГ изображения – образца; вычисление среднеквадратичного отклонения элементов ЦГ; сравнение построенной ЦГ с гистограммами изображений, хранящихся в БД, путем вычисления коэффициента корреляции  $d_i$ ; сортировка полученных значений  $d_i$ ; визуализация изображений, наиболее близких к искомому. Поэтапность и независимость обработки данных при выполнении основных шагов контекстного поиска определяют конвейерную организацию вычислений в процессе решения. Построенная на этих принципах вычислительная система может иметь структуру, приведенную на рис. 1.

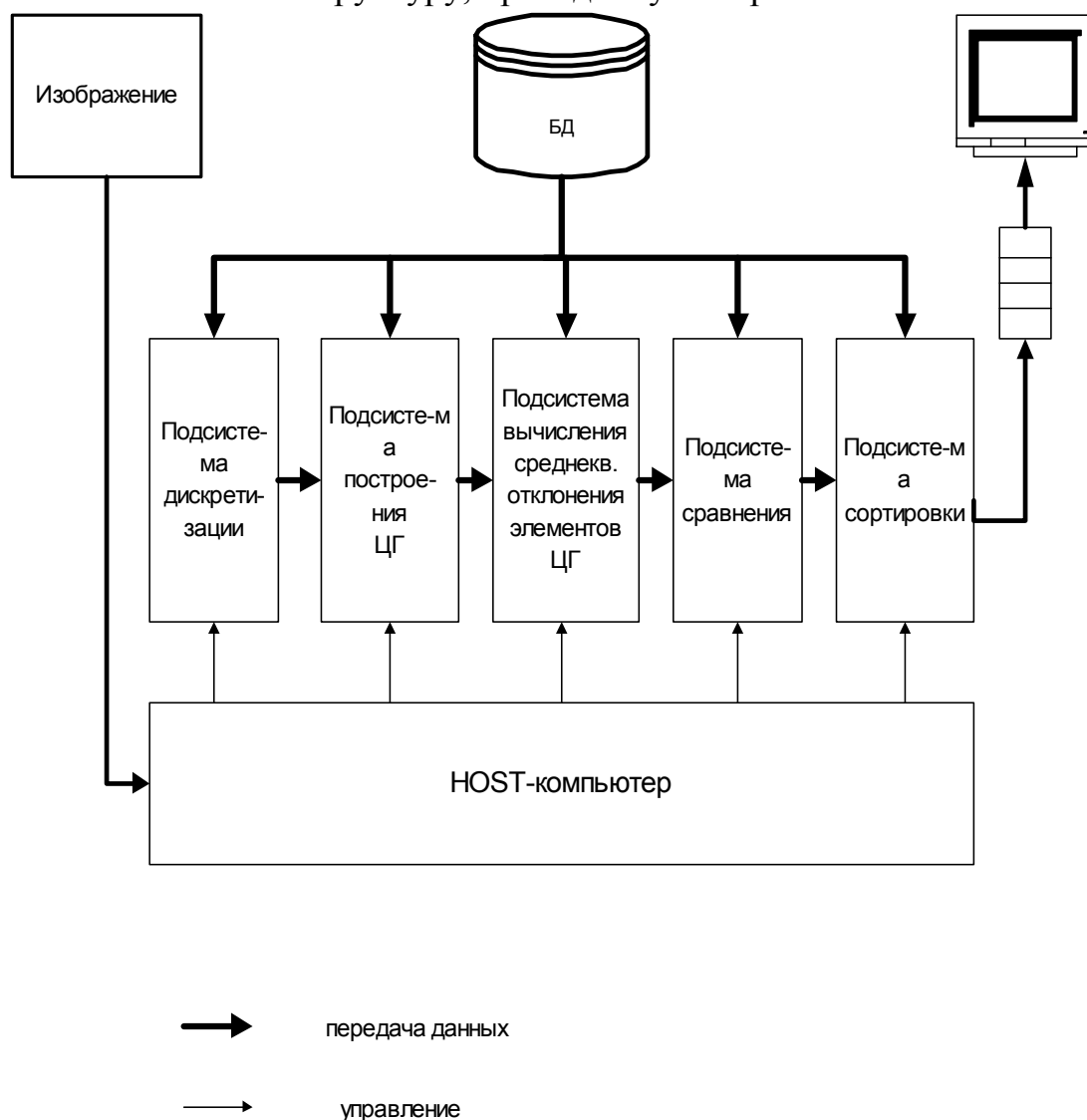


Рис. 1. Вычислительная система для контекстного поиска изображений

## 2. Исследование эффективности функционирования специализированной вычислительной системы

Вычислительная система работает под управлением HOST-компьютера, который задает последовательность работы ступеней конвейера. Каждая подсистема может быть реализована в виде SIMD-набора параллельно работающих специализированных процессорных элементов (ПЭ) упрощенной структуры, ориентированных на выполнение определенного алгоритма. Для оценки целесообразности применения какой-либо определенной топологии анализировались зависимости времени выполнения каждого этапа от размеров изображения и объема БД для каждого из этапов на различных архитектурах. Графики таких зависимостей приведены на рис. 2- 5.

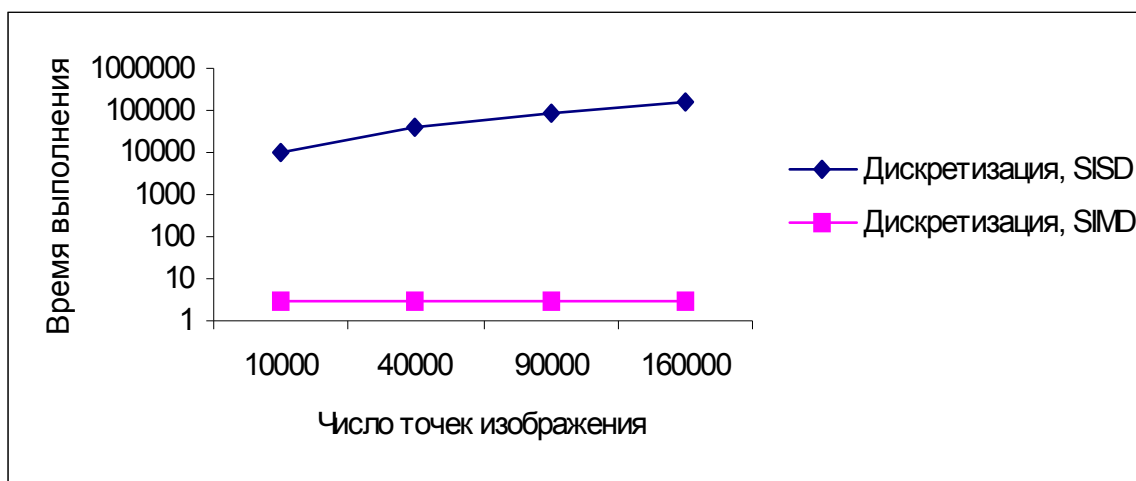


Рис. 2. Зависимость времени выполнения этапа приведения цветов изображения к базовому набору цветов на различных архитектурах от количества точек изображения

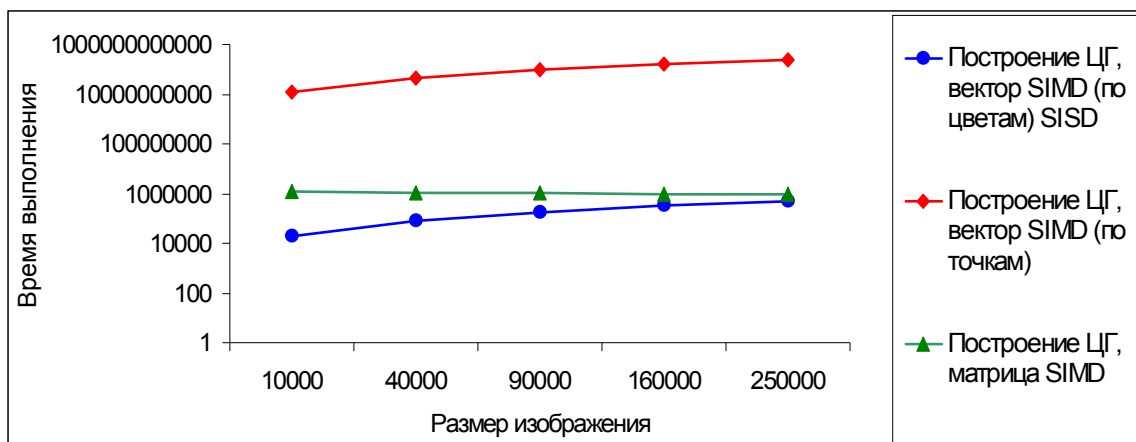


Рис. 3. Зависимость времени построения цветовой гистограммы на различных архитектурах от количества точек изображения

Проведенная оценка времени выполнения алгоритмов отдельными ступенями конвейера показала, что применение распараллеливания на всех этапах вычислительного процесса контекстного поиска изображений позволяет значительно сократить время вычислений. Была проанализирована эффективность использования для каждой из ступеней вектора либо матрицы SIMD- процессоров (для подсистемы сравнения цветовых гистограмм также исследовалась возможность использования кубической сетки). Относительно выбора архитектуры для реализации каждой ступени конвейерной вычислительной системы можно сказать следующее. Для реализации подсистемы дискретизации цветов изображения следует выбрать линейку либо матрицу процессоров (топология не имеет значения, поскольку алгоритм не предполагает использование межпроцессорного обмена данными). Для этапа построения цветовых гистограмм наилучшие временные характеристики присущи вектору процессоров, каждый из которых вычисляет отдельный элемент цветовой гистограммы, также следует использовать векторную топологию и при вычислении среднеквадратичного отклонения элементов цветовой гистограммы. Для этапа же сравнения изображений практически безразлично, какую из топологий (векторную или кубическую) использовать, поскольку их временные характеристики отличаются мало, однако из-за более простой организации, возможно, более целесообразно использовать вектор процессорных элементов.

Для каждой ступени предложенной конвейерной ВС были сформированы топология межпроцессорных связей и структуры процессорных элементов, ориентированные на выполнение конкретных алгоритмов и обеспечивающие наилучшие временные характеристики функционирования каждой из подсистем в отдельности. Далее был выполнен анализ эффективности работы вычислительной системы в целом. Главным критерием при этом являлось существование установившегося режима функционирования вычислительной системы (то есть очередь на входе любой из подсистем не должна расти бесконечно).

В ходе анализа были определены значения коэффициентов виртуализации для каждой из подсистем, обеспечивающие выполнение выдвинутого условия. При этом выяснилось, что эффективное взаимодействие подсистем конвейера, предложенного ранее, невозможно из-за его несбалансированности. Была выполнена реорганизация вычислительной системы, обеспечивающая отсутствие очереди на входе каждой из подсистем. Так, были объединены подсистемы дискретизации, построения цветовой гистограммы и подсистема вычисления среднеквадратичного отклонения элементов цветовой гистограммы в одну (далее- подсистема индексирования изображения).

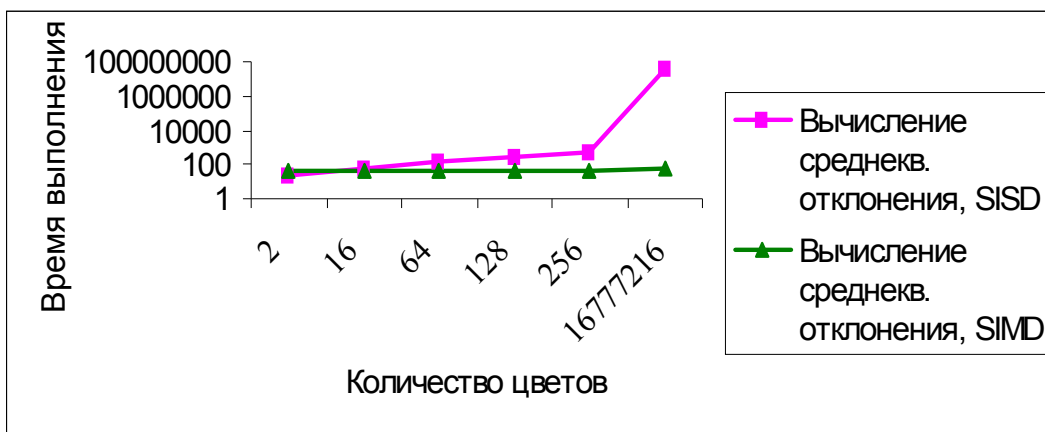


Рис. 4. Зависимость времени вычисления среднеквадратичного отклонения элементов цветовой гистограммы на различных архитектурах от количества цветов изображения

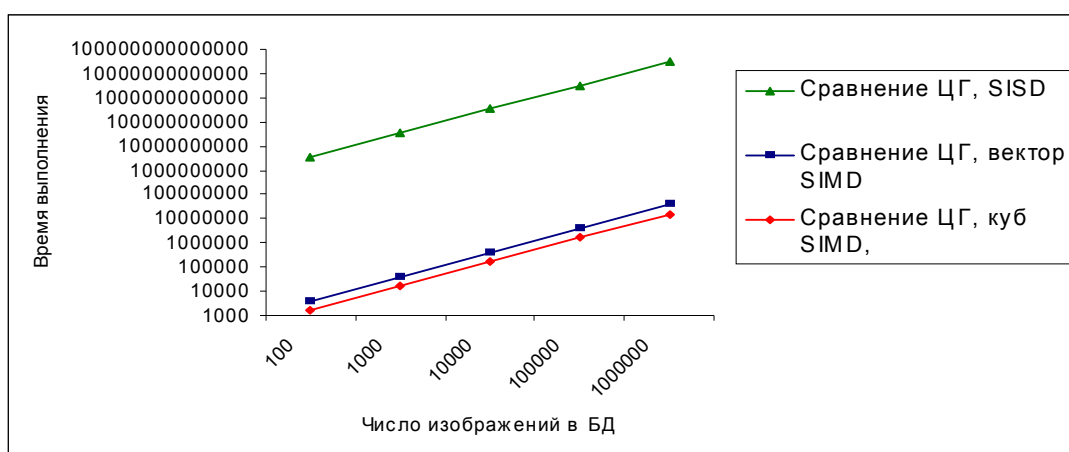


Рис. 5. Зависимость времени выполнения этапа сравнения изображений на различных архитектурах от числа изображений в БД

Данная реорганизация привела к необходимости внесения изменений в топологию межпроцессорных связей и структуру процессорных элементов. Для построения цветовой гистограммы изображения и вычисления среднеквадратичного отклонения ее элементов, как и ранее, используется линейка процессорных элементов. Аналитически была доказана эффективность работы такого конвейера.

Для проверки правильности сделанных выводов была разработана программная модель, позволяющая имитировать процесс выполнения нескольких следующих друг за другом запросов на поиск изображений в базе данных. В данном случае имитационная модель позволяет получить информацию о загрузке и времени работы процессорных элементов каждой из подсистем и построить временные диаграммы работы вычислительной системы. В ходе имитационного моделирования исследовалась работа ВС для контекстного поиска изображений до и после описанной выше реорганизации. В ходе опытов случайным образом

генерировались такие параметры работы системы, как размеры изображения- образца и объем БД изображений. Результатом моделирования является время работы каждой из подсистем конвейера до и после реорганизации. Для оценки временной сложности алгоритма был выбран базовый процессор – аналог Intel x86, и время выполнения алгоритмов оценивалось в количестве тактов этого процессора, исходя из длительности выполнения им каждой операции [7], данные о которых приведены в таблице 1.

Таблица 1

Количество тактов выполнения основных арифметических операций  
базовым процессором

Операция	Количество тактов выполнения операции с плавающей запятой	
	Одинарная точность	Двойная точность
Сложение/вычитание	1	1
Сравнение	1	1
Умножение	1	1
Деление	13	17
SQRT	16	24

По результатам моделирования с различными параметрами определялось среднее значение модельного времени работы каждой ступени, и по этим значениям были построены временные диаграммы работы вычислительной системы, приведенные на рис. 6 и 7. Из первой временной диаграммы (рис. 6) видно, что время выполнения ступеней конвейера существенно различается, что приводит к росту времени обслуживания запросов на поиск. Иначе говоря, если имеется несколько запросов на поиск изображений в БД, то каждый следующий запрос будет выполняться дольше предыдущего. Данная ситуация обусловлена несбалансированностью предложенного конвейера. Временная диаграмма работы реорганизованного конвейера, состоящего из двух подсистем, приведенная на рис. 7, показывает, что в результате реорганизации время обработки запроса на поиск стало величиной постоянной. Это подтверждается построенными по результатам моделирования графиками зависимости времени выполнения от количества поступивших запросов на поиск (рис.8). Как видно из рисунка, до реорганизации вычислительной системы время выполнения очередного запроса превышало время выполнения предыдущего. Выполненная реорганизация позволила достичь того, что, при наличии нескольких следующих друг за другом запросов время ответа системы остается постоянным, а не увеличивается с поступлением каждого нового запроса, то есть, было обеспечено корректное функционирование вычислительной системы в установившемся режиме.

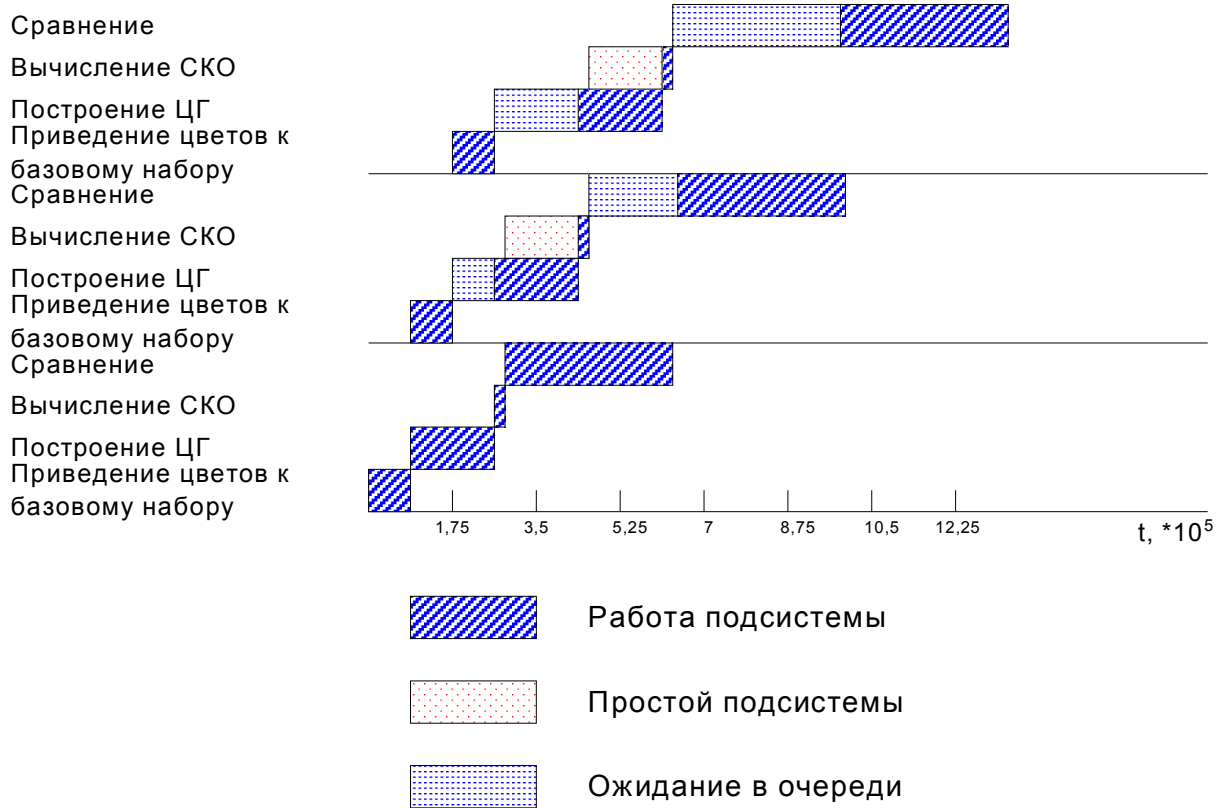


Рис. 6. Временная диаграмма работы конвейера для контекстного поиска изображений до реорганизации

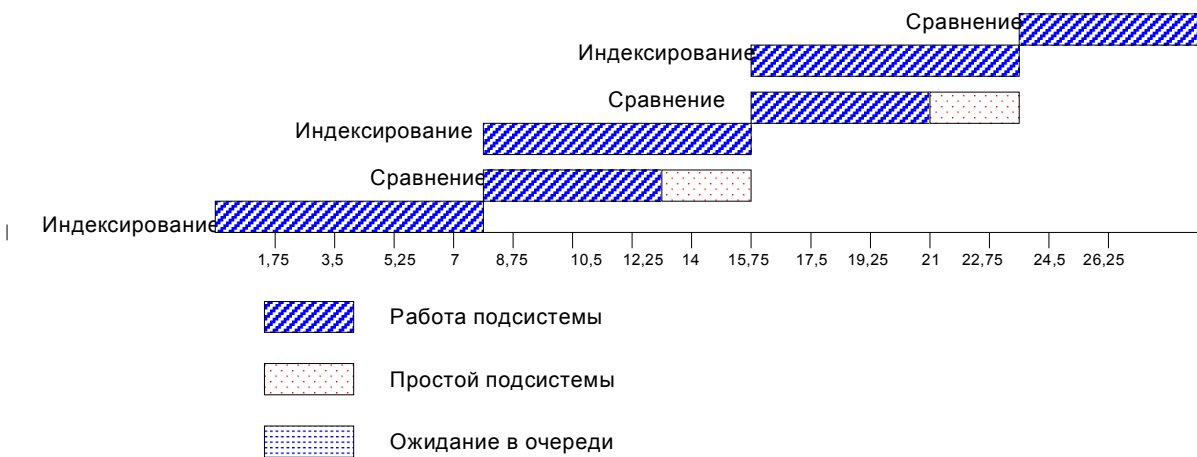


Рис. 7. Временная диаграмма работы вычислительной системы для поиска изображений после реорганизации

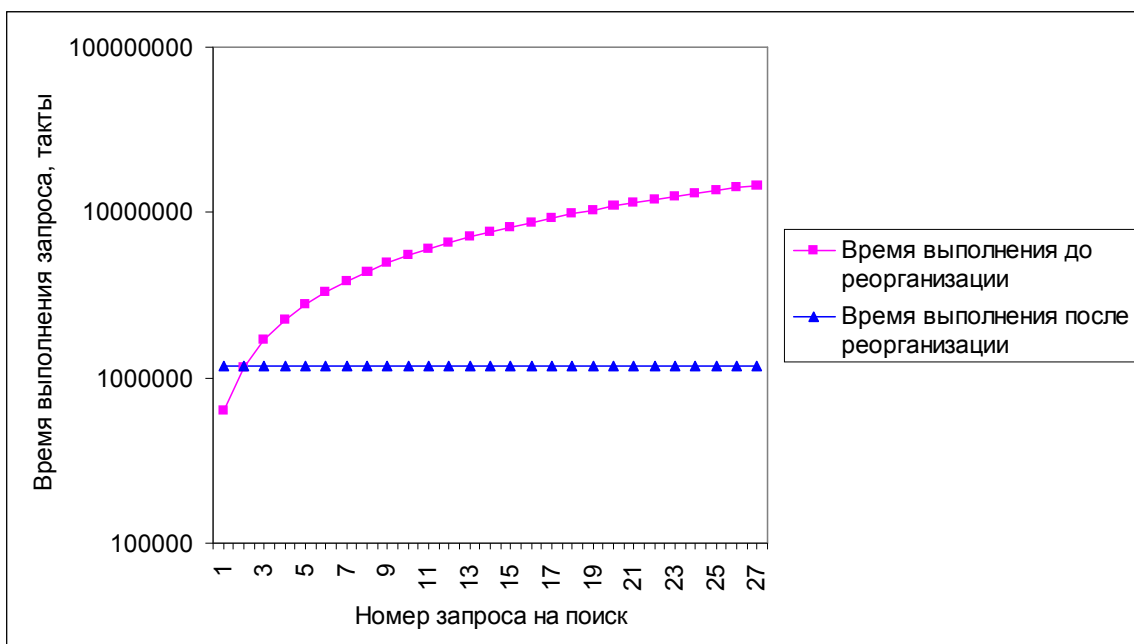


Рис. 8. Зависимость времени выполнения запроса на поиск от количества последовательно сформированных запросов

## Выводы

Для решения задачи поиска изображений на основе их цветового содержимого характерна поэтапность и независимость этапов обработки, что позволило выбрать для вычислительной системы в целом конвейерную архитектуру. Проведенная оценка времени выполнения алгоритмов отдельными ступенями конвейера показала, что применение распараллеливания на всех этапах вычислительного процесса контекстного поиска изображений позволяет значительно сократить время вычислений на каждом этапе. Однако сформированные для каждой ступени первоначально предложенной конвейерной ВС топология межпроцессорных связей и структуры процессорных элементов, ориентированные на выполнение конкретных алгоритмов и обеспечивающие наилучшие временные характеристики функционирования каждой из подсистем в отдельности, приводят к несбалансированности конвейера в целом. В связи с этим предложена и аналитически обоснована его реорганизация, обеспечивающая отсутствие очереди на входе каждой из подсистем, показана эффективность выполненной реорганизации.

В дальнейшем предполагается проанализировать применимость изложенных идей для поиска изображений по текстурным признакам.



## **Литература**

1. Simon D. Haynes, John Stone, Peter Y.K. Cheung, Wayne Luk. Video Image Processing with the Sonic Architecture// IEEE Computer.–Aprils, 2000.– pp.50–57.
2. R.M.Lea. A VLSI array processor for image processing// Algorithmically specialized parallel computers. Edited by L. Snyder, L.H.Jamieson, D.B.Gannon, H.J.Siegel. Academic Press,Inc. 1985. pp.159-168
3. Башков Е.А., Шозда Н.С. Использование специализированной ЭВМ для решения задачи контекстного поиска изображений. Наукові праці Донецького державного технічного університету. Серія: "Проблеми моделювання та автоматизації проектування динамічних систем". Випуск 10. Донецьк, 1999. - с. 247-252.
4. Шозда Н.С. Ускорение вычислений на этапе сравнения цветowych гистограмм в процессе контекстного поиска изображений. Наукові праці Донецького державного технічного університету. Серія: "Проблеми моделювання та автоматизації проектування динамічних систем". Випуск 29. Севастополь –Донецьк, 2001. с. 129-134
5. Шозда Н.С. Модифицированный алгоритм поиска в базе данных изображений на основе их цветового содержимого. III Міжнародна молодіжна науково-практична конференція "Людина і космос", присвячена 40-річчю першого польоту людини в космос: Збірник тез. Дніпропетровськ, НЦАОМУ, 2001 - с. 298.
6. Башков Е.А., Шозда Н.С. Поиск изображений в больших БД с использованием коэффициента корреляции цветowych гистограмм.// Труды 12-й Международной конференции ГрафиКон'2002. – Нижний Новгород, 2002. – стр. 358–361. 12-я Международная конференция по компьютерной графике и машинному зрению ГрафиКон'2002. Нижний Новгород (Россия), 16-21 сентября 2002 г.
7. Кузьминский М. Athlon: от микропроцессоров к материнским платам// Открытые системы. – 2000. – № 1–2. – с. 8-13.

---

Дата надходження до редколегії: 11.11.2003 р.