

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

На правах рукописи

БЛИЗНЮК Алексей Владимирович

**РАЗРАБОТКА И РЕАЛИЗАЦИЯ WEB-СЕРВИСОВ ДЛЯ СРАВНЕНИЯ И
ОБРАБОТКИ
ВИДЕОИЗОБРАЖЕНИЙ**

05.13.11 – Математическое и программное обеспечение
вычислительных машин, комплексов и компьютерных сетей

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата физико-математических наук

Санкт-Петербург

2010

Работа выполнена на кафедре информатики математико-механического факультета Санкт-Петербургского государственного университета.

Научный руководитель:	доктор технических наук, профессор САФОНОВ Владимир Олегович
Официальные оппоненты:	доктор технических наук, профессор ЛИСС Александр Рудольфович (ОАО "Концерн "Океанприбор") кандидат физико-математических наук, доцент ФОМИНЫХ Николай Федорович (ЗАО «Ланит-Терком»)
Ведущая организация:	Санкт-Петербургский государственный политехнический университет

Защита состоится «___» _____ 2010 года в _____ часов на заседании совета Д 212.232.51 по защите докторских и кандидатских диссертаций при Санкт-Петербургском государственном университете по адресу: 198504, Санкт-Петербург, Петродворец, Университетский пр., д. 28, математико-механический факультет Санкт-Петербургского государственного университета, ауд. 405.

С диссертацией можно ознакомиться в Научной библиотеке Санкт-Петербургского государственного университета по адресу: 199034, Санкт-Петербург, Университетская наб., д. 7/9

Автореферат разослан «___» _____ 2010 года.

Ученый секретарь
диссертационного совета,
доктор физико-математических наук,
профессор

И. К. Даугавет

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Данная работа посвящена разработке методов обработки и хранения больших объемов видеoinформации с помощью облачных вычислений (cloud computing), а также реализации программного комплекса, основанного на этих методах.

Актуальность темы. Задачи, связанные с обработкой видеoinформации, привлекают к себе все больше внимания, в связи со все более широким распространением систем видеонаблюдения и регистрации в различных сферах человеческой деятельности. Данная диссертационная работа посвящена разработке методов обработки видеoinформации с помощью облачных вычислений, основанных на распределенной сервисно-ориентированной архитектуре и позволяющих более эффективно использовать мощности вычислительных центров. При разработке методов предложены решения проблем передачи и хранения больших объемов информации, а также разработаны эффективные методы обработки изображений, устойчивые к помехам. В рамках работы реализован программный комплекс, представляющий собой набор web-сервисов, обеспечивающих решение задач обработки видеоизображений с помощью облачных вычислений в сети Интернет [1].

Цели работы. Цели диссертационной работы заключаются в разработке методов обработки и хранения видеoinформации с помощью облачных вычислений, а также в реализации распределенного программно-аппаратного комплекса, основанного на сервисно-ориентированной архитектуре, позволяющего максимально эффективно использовать вычислительные ресурсы системы, балансируя нагрузку между клиентскими узлами и фермой серверов.

Для достижения указанных целей в работе решены следующие задачи:

- Разработаны методы обработки и хранения видеoinформации с помощью облачных вычислений
- Разработан и реализован быстрый алгоритм сравнения двух графических изображений с учетом помех
- Разработан набор Web-сервисов для обработки, передачи и хранения видеоизображений
- Разработана и реализована библиотека функций для работы с изображениями: сравнения изображений различными методами, изменения яркости, изменения размеров
- Разработан и реализован пользовательский интерфейс для управления разработанной системой обработки изображений

Результаты решения этих задач выносятся на защиту.

Предметом исследования являются методы работы с видеоизображениями, хранения и передачи больших объемов данных.

Методы исследования. Для исследования возможностей системы был использован метод моделирования отдельных частей системы и применения к ним эмпирических метрик, что дало возможность установить на реальных экспериментах возможности и технические характеристики будущей системы.

Обоснованность и достоверность программного обеспечения получены путем тестирования и использования программно-аппаратного комплекса на реальных объектах.

Научные результаты и их новизна. В ходе диссертационной работы разработаны web-сервисы для работы с видеоизображением. Суть работы сводится к следующему:

- предложен набор методов обработки и хранения видеоинформации в реальном времени
- создан программно-аппаратный комплекс, представляющий собой систему видеонаблюдения, не имеющий аналогов. Получен патент на полезную модель из федерального института промышленной собственности Российской Федерации № 90230, «Система видеонаблюдения с общим сервером в сети Интернет», автор Близняк А.В., заявка № 2009113530, на 8 страницах, включая описание, реферат, формулу полезной модели и схемы;
- разработаны способы передачи больших объемов информации без потери производительности;
- реализованы методы хранения видеоизображения на сервере системы;
- проанализированы и реализованы методы обработки видеоизображений, в результате чего получена библиотека, позволяющая максимально быстро сравнивать изображения с учетом помех;
- разработан новый алгоритм для поиска различий на изображениях;

Практическая значимость работы. Результатом работы является программно-аппаратный комплекс для видеонаблюдения и обработки видеоизображений, позволяющий конечному пользователю значительно сократить свои затраты, а также имеющий ряд конкурентных преимуществ перед ближайшими аналогами. Разработанный программно-аппаратный комплекс внедрен на предприятии «Ланит-Терком» и получил положительные отзывы пользователей.

Апробация работы. Основные выводы и результаты докладывались на конференции «Технологии Microsoft в теории и практике программирования» [3], где автором было занято второе место и на конференции «Первый конкурс ИТ проектов», где

автор занял первое место. В 2009 году был выигран грант Microsoft Старт 09. По материалам диссертации опубликованы 3 статьи, в том числе статьи из списка ВАК.

Публикации. Основные результаты диссертации опубликованы в работах [1-4]. Из них две публикации [1, 2] в журналах из перечня ВАК. Работа [3] написана в соавторстве: диссертанту принадлежат общие постановки задач, реализация описываемых методов, создание программных средств, соавтору – основные принципы реализации распределенной системы с помощью web-сервисов.

Структура и объем диссертации. Работа содержит введение, три главы и заключение. Объем работы 98 страниц, количество иллюстраций 8, список использованной литературы содержит 32 наименования.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Введение содержит предварительную информацию о предмете исследования, список некоторых задач, решаемых с использованием методов, описанных в диссертации, а также формулировки целей диссертации.

В **первой главе** приводится обзор существующих схем построения видеонаблюдения. Дается краткое изложение особенностей архитектуры и применения множества современных систем обработки и хранения изображений. Описаны различные способы обработки изображений. В *попиксельном методе* сравнения изображение представляется в виде матрицы размера $n \times m$ со значениями каждого элемента от 0 до 255. Пусть матрица A – исходное изображение, матрица B – второе изображение. Попиксельный метод сравнения заключается в том, что для каждого элемента ij ($i=1..n$, $j=1..m$) будем искать разность $R_{ij} = |a_{ij} - b_{ij}|$ и накапливать количество значений пикселей, для которых R отлично от 0 (то есть, пиксели различаются).

$$T_{ij} = \begin{cases} 0, R_{ij} = 0 \\ 1, R_{ij} \neq 0 \end{cases}$$

$$T = \frac{1}{nm} \sum_{i,j} T_{ij}$$

Введем порог ошибки E – допустимая доля отличающихся пикселей в изображениях.

Если $T \leq E$, то принимается вариант, что изображения одинаковы, если $T > E$ – изображения различны.

Статистический подход основывается на предположении, что исходное пространство объектов представляет собой вероятностное пространство, а признаки (характеристики) объектов являют собой случайные величины, заданные на нем. Будем рассматривать набор пикселей как статистическую выборку. Тогда для

определения подобности изображений будем использовать статистические методы сравнения выборок.

Критерий Пирсона или критерий χ^2 (хи - квадрат) имеет наибольшее применение при проверке согласования теоретической и эмпирических кривых распределения. Наблюдаемое значение критерия $(Y)H = \chi^2$ вычисляется по следующей формуле:

$$\chi^2_H = \sum_{i=1}^n \frac{(m_{эi} - m_{тi})^2}{m_{тi}},$$

где $m_{эi}$ - эмпирическая частота i -го интервала (варианта);

$m_{тi}$ - теоретическая частота i -го интервала (варианта);

n - число интервалов (вариантов).

Как известно χ^2 - распределение зависит от числа степеней свободы, это число находится по формуле

$$v = n - r - 1,$$

где r - число параметров предполагаемого теоретического закона, использованных для вычисления теоретических частот и оцениваемых по выборке.

Коэффициент ранговой корреляции Спирмена – это непараметрический метод, который используется с целью статистического изучения связи между явлениями. Практический расчет коэффициента ранговой корреляции Спирмена включает следующие этапы:

- 1) Сопоставить каждому из признаков их порядковый номер (ранг) по возрастанию (или убыванию).
- 2) Определить разности рангов каждой пары сопоставляемых значений.
- 3) Возвести в квадрат каждую разность и суммировать полученные результаты.
- 4) Вычислить коэффициент корреляции рангов по формуле:.

$$r = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

где $\sum d^2$ - сумма квадратов разностей рангов, а n - число парных наблюдений.

Во **второй главе** детально рассматривается система видеонаблюдения с общим сервером в сети Интернет, концепции и принципы, положенные в основу разработки системы. Описана архитектура системы, механизмы работы ее компонент и взаимодействие между ними [1]. Приводится обзор основных технологий, использованных при создании

комплекса. В главе рассмотрен новый подход к построению систем видеонаблюдения. Инновационной особенностью системы является то, что серверный кластер расположен в сети Интернет, к нему подсоединяются через каналы Интернета видеокамеры, с которых данные идут на сервер. Клиент имеет возможность просматривать записанные данные с помощью любого компьютера, подсоединенного к Интернету (рисунок 1).

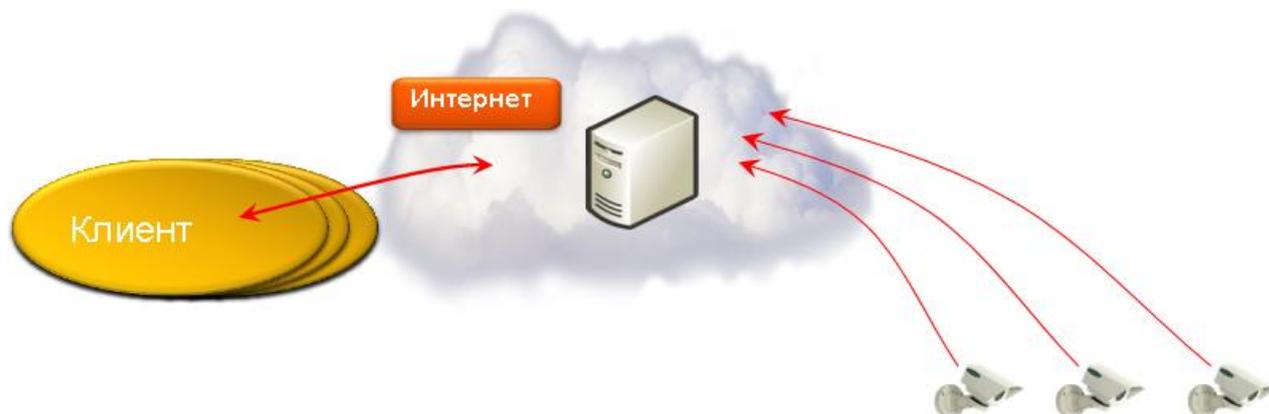


Рисунок 1 Система видеонаблюдения с общим сервером в сети Интернет

Во второй части главы рассматриваются теоретические основы применения СОМ-объектов для обеспечения независимости от платформы [2]. Хорошо известны способы переиспользования кода, описывающего некоторую функциональность в рамках одного языка программирования, одной платформы, одного компилятора. Использование же модуля, скомпилированного из исходного кода, написанного на одном языке, в программном коде, написанном на другом языке, было затруднено. Для решения данной задачи компания Microsoft предложила технологию СОМ (Component Object Model) – компонентную объектную модель. Компонент – это независимая часть программного обеспечения, предназначенная для выполнения какой-то определенной работы. Функциональность компонента может быть использована в любой программе. Нет необходимости дублировать код по разным программам, в которых требуется использовать одну и ту же функциональность.

В конце главы приведены варианты сравнения изображений с помощью спектрального анализа, описан метод быстрого преобразования Фурье для получения спектра изображения.

Третья глава полностью посвящена деталям реализации соискателем системы видеонаблюдения, основным частям инфраструктуры, ресурсоемкости, выбранным технологиям, подсистеме, отвечающей за взаимодействие с пользователем.

В главе подробно описывается реализация пользовательского интерфейса, модуля, отвечающего за лицензирование, способа хранения записей видеоизображений, сохранения конфигурации. Рассматривается общая архитектура комплекса.

Для реализации пользовательского интерфейса применялась технология Windows Presentation Foundation (рисунок 2). WPF предоставляет разработчикам объединенную

модель для программирования пользовательских интерфейсов, графических объектов и документов. Пользовательский интерфейс прототипа был разделен на две части: область отображения данных с видеокамер и область для управления камерами. В области отображения видеокамер может быть несколько окон, автоматически выполняющих выравнивание в области.



Рисунок 2 Пользовательский интерфейс

Для задач лицензирования был использован RSA-алгоритм. RSA (аббревиатура фамилий Rivest, Shamir, Adleman) — криптографический алгоритм с открытым ключом.

Для хранения записей была разработана следующая структура хранения данных. Изображения хранятся в виде двух файлов: бинарный файл и файл с информацией для получения данных в формате .xml .

Запись происходит следующим образом:

- изображение – это набор кадров.
- каждый кадр конвертируется в набор байтов.
- Эти наборы байтов последовательно записываются в бинарный файл, а необходимая для чтения информация (метаданные): число байтов в кадре, смещение относительно начала файла, точное время кадра и значение частоты кадров записываются в отдельный файл метаданных.

На рабочей машине оператора сохраняется информация о том, с каких камер велась запись, какие камеры просматривались на момент выхода из программы, сохраняется конфигурация для каждой камеры, информация о локализации программного продукта,

добавлен ли этот программный продукт в автозапуск системы Windows. Информация о конфигурации видеокамер хранится в файле формата xml "devices.xml".

Информация о том, с каких камер велась запись видеопотока на компьютер, и информация о том, какие камеры были активны в момент выхода из данного программного продукта, хранится в файле "workers.xml".

В последней части главы приводится описание реализации предложенного метода поиска различий на изображениях с помощью спектрального анализа и доказывается устойчивость к помехам при применении этого подхода (рисунок 3).

В обработке изображений преобразование Фурье обычно рассматривается как декомпозиция сигнала на частоты и амплитуды, то есть, обратимый переход от временного пространства (time domain) в частотное пространство (frequency domain). В частотном пространстве каждая точка описывает соответствующую частоту из временного пространства. Преобразование Фурье может быть применено для избавления от помех на изображении. Поскольку речь идет о цифровых изображениях, будет применяться дискретное преобразование Фурье (ДПФ). Дискретное преобразование будет содержать не все частоты, формирующие изображение, а только те, которые достаточно велики, чтобы описать временное пространство изображения. Количество частот соответствует количеству пикселей исходного пространства изображения, таким образом изображения во временном и частотном пространствах имеют одинаковые размерности.

Для квадратного изображения размером $N \times N$, ДПФ представлено следующей формулой:

$$F(k, l) = \frac{1}{N^2} \sum_{a=0}^{N-1} \sum_{b=0}^{N-1} f(a, b) e^{-i2\pi(\frac{ka}{N} - \frac{lb}{N})}$$

Где $f(a,b)$ – изображение во временном пространстве и экспонента – базисная функция, соответствующая каждой точке $F(k,l)$ в частотном пространстве. Таким образом, значение каждой точки $F(k,l)$ вычисляется перемножением пространственного изображения с соответствующей базисной функцией и суммированием результатов. Базисные функции представляют собой синусы и косинусы с увеличивающимися частотами, таким образом $F(0,0)$ представляет главный коэффициент, который характеризует среднюю яркость изображения, а $F(N-1,N-1)$ представляет самую высокую частоту.

Вычислительная сложность ДПФ составляет N^2 . Для уменьшения сложности до $N \log_2 N$ используем быстрое преобразование Фурье. Это улучшение особенно заметно для изображений больших размеров. Для применения БПФ изображение должно быть приведено к размеру 2^N . При этом полезно произвести уменьшение размера изображения, что увеличит скорость вычислений и не сильно скажется на качестве.

Преобразование Фурье переводит изображение в пространство, состоящее из комплексных чисел, которое может быть представлено двумя изображениями, отвечающими за вещественные и мнимые части или амплитудные и частотные

характеристики. Для обработки изображения наиболее важна вещественная часть, т.к. она содержит большую часть информации об его геометрической структуре.

Изучая частотные характеристики, можно понять, как они влияют на изображение во временном пространстве. Обычно для удобного отображения на экране частотное пространство сдвигается так, чтобы главный коэффициент оказался в середине. Чем дальше точка расположена от центра, тем выше ее частота.

Получив изображение частотного пространства, видно, что наибольшее значение у главного коэффициента, который расположен в центре, значения остальных коэффициентов слишком велики, поэтому остальная часть окрашена в черный цвет.

Для получения более интересного изображения применим к каждому коэффициенту логарифмический оператор:

$$c = \frac{255}{\log(1 + |R|)}$$

Полученное изображение содержит значения всех частот, где, чем выше частота, тем меньше ее амплитудное значение. Таким образом, низкие частоты содержат больше полезной информации, чем высокие. Это означает, что небольшая помеха на исходном изображении слабо скажется на основных амплитудных значениях. Отбрасываем высокие частоты и изучаем оставшиеся амплитуды двух изображений:

$$S = \frac{\sum_{i=\frac{N}{2}-d}^{\frac{N}{2}+d} \sum_{j=\frac{N}{2}-d}^{\frac{N}{2}+d} F(i,j)}{4d^2},$$

где значение d устанавливается эмпирически.

Сложность метода составляет $O(N \log_2 N) + 4d^2$.

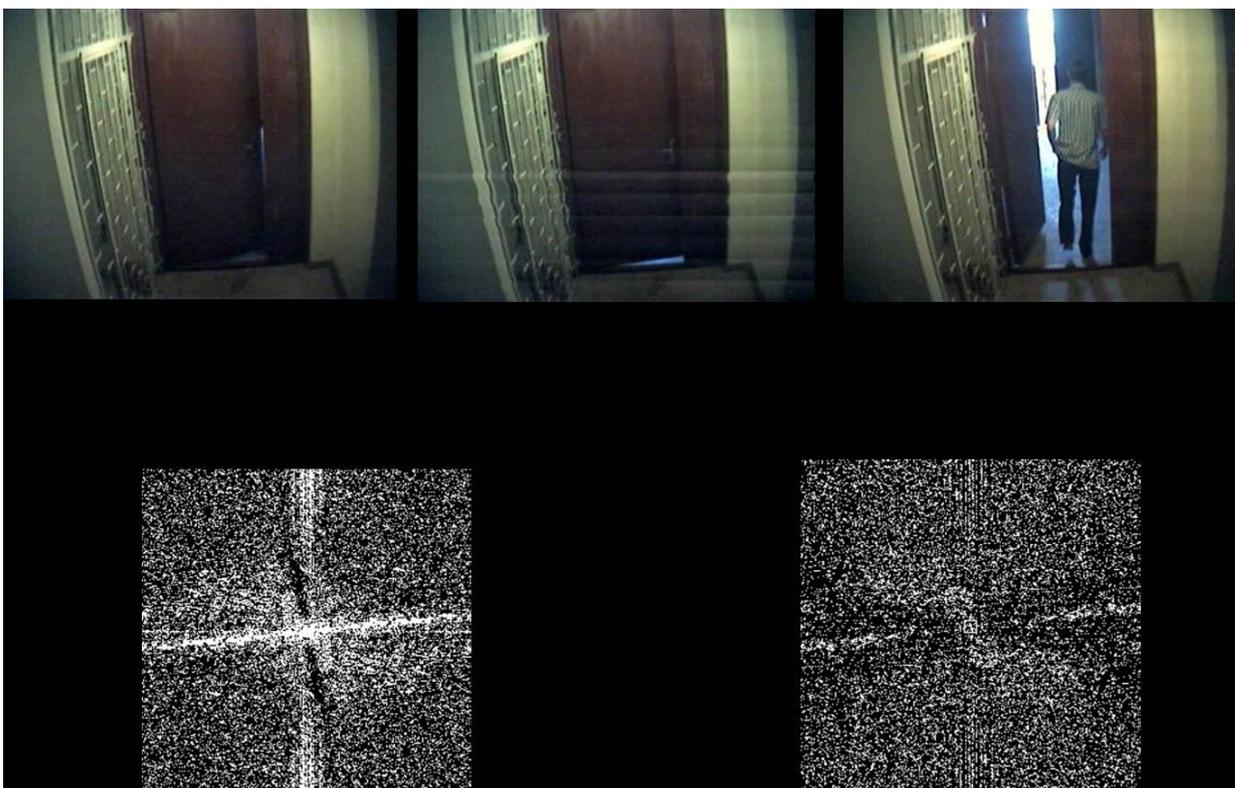


Рисунок 3 Пример изображений и их спектров

Разработанная как практический результат диссертации библиотека, включающая в себя реализацию приведенного метода сравнения изображений, является продуктом с открытым кодом и доступна на сайте [google code](https://github.com). Для сравнения эффективности метода проанализируем аналогичные наиболее известные библиотеки с открытым кодом: AForge.NET, Python Motion Detection Library, Soft Collection Motion Detection Library.

Библиотека AForge.NET, реализованная на языке C#, предназначена для разработчиков и исследователей в области компьютерного зрения и искусственного интеллекта - нейронных сетей, генетических алгоритмов, робототехники и т. д., и содержит реализацию алгоритмов обработки изображений. В пространстве имен AForge.Vision.Motion имеется класс MotionDetector с методом ProcessFrame, позволяющим сравнивать два изображения. Метод основан на одной из модификаций попиксельного сравнения. Сложность такого алгоритма $O(N^2)$.

	Image Processing (диссертация)	AForge.NET	Python Motion Detection Library	Soft Collection Motion Detection Library
Сложность алгоритма	$O(n \log_2 n) + 4d^2$	$O(n^2)$	$O(n^2 + 8 * n/9)$	$O(n^2)$
Помехоустойчивость	есть	нет	есть	нет
Использование облачных вычислений (cloud computing)	есть	нет	нет	нет

Таблица 1. Сравнение разработанной библиотеки функций с аналогами.

В таблице 1 приведены сравнительные характеристики реализованной библиотеки для сравнения изображений с наиболее близкими аналогами.

Заключение содержит описание основных результатов, полученных в работе, а также перечень возможных направлений развития проекта.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ И ВЫВОДЫ

В ходе выполнения работы достигнуты следующие результаты:

1. Разработан программный комплекс оригинальной архитектуры для видеонаблюдения на основе применения концепции масштабирования физических ресурсов («облачные вычисления»), с сервером в сети Интернет.
2. Разработан быстрый алгоритм (таблица 1) для сравнения изображений с помехами, выполнена оценка его сложности, разработан реализующий его программный модуль.
3. Разработан набор web-сервисов для обработки, передачи и хранения видеоизображений.

Все полученные результаты являются новыми.

Реализованный программно-аппаратный комплекс успешно прошел испытания и эксплуатируется на многих объектах.

Предложенный вариант реализации системы видеонаблюдения, где единый сервер хранения информации располагается в сети Интернет, выделил целый ряд преимуществ такой системы перед теми, которые уже существуют:

- Увеличение производительности, благодаря использованию облачных вычислений

- Значительное снижение требований к ресурсам на локальной машине, в связи с тем, что пользователь получает необходимые ресурсы по мере необходимости. Благодаря предлагаемой архитектуре системы, ее администрирование является задачей поставщиков web-сервисов.
- Независимость местонахождения и платформная независимость достигается за счет использования сервисов с помощью web-браузера на любом устройстве (ПК, мобильный телефон, карманный компьютер). Поскольку сервисы доступны через Интернет, пользователи могут работать с ними с любого месторасположения, где есть доступ в сеть.
- Централизация хранилища данных пользователей решает проблемы разделения ресурсов между пользователями, улучшает распределение нагрузки в силу того, что большинство пользователей загружают систему не более чем на 10-20%.
- Повышается надежность сервера, что предотвращает возможность потери или порчи данных.
- Улучшение масштабируемости за счет динамического распределения ресурсов между пользователями, самообслуживания сервера в режиме реального времени. Решение различных слабо связанных задач представлено в виде различных web-сервисов, что является наиболее важным и новым методом в преодолении узких мест производительности, позволяя равномерно и по мере необходимости распределять ресурсы системы.
- Благодаря централизации хранилища, становится возможным значительно усилить защиту данных, используя защищенные каналы связи и авторизацию пользователей.

В ходе выполнения работы был получен патент на полезную модель, описывающий основные моменты созданного программного комплекса [4]. На конкурсе фонда Бортника проект занял второе место по Северо-западному федеральному округу, были проведены НИОКР в соответствии с календарным планом государственного контракта № 6847р/9563. Апробация системы на промышленных и социальных объектах позволила объективно оценить эффективность предложенного соискателем решения.

РАБОТЫ АВТОРА ПО ТЕМЕ ДИССЕРТАЦИИ

Статьи в журналах, рекомендованных ВАК:

- [1] Близнюк А.В. Система видеонаблюдения с общим сервером в сети Интернет. // Научно-технические ведомости СПбГПУ. 2009, Вып. 4, С. 141-147
- [2] Близнюк А.В. Создание и применение компонентов COM. // Вестник Санкт-Петербургского университета. 2010 (март), Сер. 10, Вып. 1, С. 117-129

Другие публикации:

- [3] Близнюк В.А., Сафонов В.О. Система видеонаблюдения с общим сервером в сети Интернет. // Тезисы докладов конкурса-конференции студентов, аспирантов и молодых ученых Северо-Запада "Технологии Microsoft в теории и практике программирования". СПб.: изд-во СПбГПУ, 2009, С. 19.
- [4] Патент на полезную модель «Система видеонаблюдения с общим сервером в сети Интернет». Автор Близнюк А. В. Дата регистрации 27 декабря 2009. Номер 90230