

# Design and implementation of a software-based vision mixer

Oier Lauzirika Zarrabeitia

March 13, 2021

**Resumen**

Lorem ipsum

**Abstract**

Lorem ipsum

# Contents

<b>Preámbulo</b>	<b>5</b>
<b>1 Introduction</b>	<b>6</b>
<b>2 State of the art</b>	<b>7</b>
<b>3 Involved technology</b>	<b>9</b>
<b>4 Description</b>	<b>10</b>
<b>5 Results</b>	<b>11</b>
<b>6 Budget</b>	<b>12</b>
<b>7 Conclusions</b>	<b>13</b>

# List of Figures

2.1 Realización del streaming de FLL 2020. Fuente: Twitter GATE . . . . . 8

# List of Tables

# Preámbulo

Lorem ipsum

# Chapter 1

## Introduction

El objetivo de este proyecto es crear un software que sea capaz de emular un mezclador de vídeo presente en cualquier instalación de producción audiovisual. Además, también se pretenden integrar otros equipos que en centros de producción tradicionales serían equipos discretos, tales como VTRs o generadores de mosaico. En cuanto a la E/S disponible, el mezclador trabajará primordialmente con fuentes de vídeo en streaming, pero permitiendo el uso de otros tipos de entrada y salida. Por último, el control se realizará por medio de comandos sobre IP, de forma que pueda ser operado de forma remota por medio de un panel físico o virtual.



## Chapter 2

# State of the art

Desde los inicios de la retransmisión televisiva a comienzos de los años 40, la realización multicámara ha sido una realidad, ya que las técnicas de narración audiovisual ya habían comenzado a madurar anteriormente con el celuloide. Por consecuencia, se desarrolló un elemento donde confluían todas las señales de cámara de los estudios de televisión, el mezclador de vídeo. Originalmente, únicamente podían conmutar entre entradas, pero con el tiempo se han desarrollado técnicas para poder realizar transiciones y composiciones entre ellas.

Hasta los años 90, estos mezcladores se basaban en circuitos analógicos, pero ya a mediados de los 90, prácticamente toda la oferta de mezcladores eran digitales. Nótese que este cambio se produce décadas antes de la transición a la televisión digital. Esto se debe a que los mezcladores digitales ofrecen una mayor flexibilidad, además de ofrecer más prestaciones en un volumen más compacto.

A mediados de la década de los 2000 se introdujeron los mezcladores basados en software. Se usan primordialmente en entornos de producción de bajo coste, ya que ofrecen altas prestaciones a costa de una menor garantía de estabilidad. Una de las primeras compañías en comercializar este tipo de software es “NewTek”, el creador del protocolo NDI anteriormente mencionado. Esta empresa vende la gama de productos “TriCaster” que consisten en ordenadores genéricos -a excepción de una tarjeta de captura de vídeo- con su software preinstalado. Como ejemplo, es la solución adoptada por el servicio audiovisual de la UPM para la realización de retransmisiones en directo. Otros productos similares a considerar pueden ser “vMix” y “Ross Graphite”, aunque esta última requiere de una tarjeta de PCIe específica para la aplicación. Aun así, marcas más tradicionales como Sony siguen sin introducirlos en su gama de productos, probablemente por su estabilidad reducida.



Figure 2.1: Realización del streaming de FLL 2020. Fuente: Twitter GATE

En el mundo open-source, la solución más madura es el Open Broadcaster Software -OBS-, ampliamente usado por “streamers” de videojuegos, aunque prácticamente es su único ámbito de uso. Por otra parte, el software “FreeJ” está orientado a la mezcla de sesiones VJ -contenidos visuales que acompañan a espectáculo de música electrónica-, por lo que prácticamente no contempla el uso de cámaras. Por consecuencia, no existe ninguna solución de código abierto orientado a la realización multicámara, por lo que éste será el objeto de este proyecto.

## Chapter 3

# Involved technology

La aplicación estará escrita en el lenguaje de programación C++. El núcleo de la aplicación yace sobre la librería Zuazo, una librería que permite manipular vídeo en tiempo real, y que es el proyecto personal del autor de este trabajo.

La librería en cuestión aprovecha en la medida de lo posible la GPU del ordenador, ya que el procesamiento del vídeo se realiza por medio de la API Vulkan, utilizando “shaders” codificados en el lenguaje GLSL. Otra de las principales características de la librería, es que se basa en clases que imitan un diagrama de bloques, lo que formalmente se conoce como “basado en grafos”. Se pueden diseñar clases que expongan una E/S arbitraria -no sólo de vídeo-, cuya implementación puede ser de bajo nivel -trabajar manualmente con fotogramas- o una composición de varias clases de complejidad inferior. Para este proyecto se tratará de utilizar mayoritariamente la segunda técnica, logrando como objetivo final clases que emulen elementos complejos de un mezclador -bancos M/E, generador de mosaico, reproductor multimedia, etc-.

Con esto se logrará implementar un modelo de software del mezclador. Sin embargo, a esta interfaz deberá de poder comunicarse con un sistema de control por red. Una de las opciones más interesantes para ello puede ser Open Sound Control -OSC- que aunque su nombre sugiere que sólo puede usarse con audio, la realidad es que se puede usar en casi cualquier ámbito. En esencia, el propósito de este protocolo es transmitir con la menor latencia posible parámetros que reflejen el estado de un panel de control físico o virtual, de forma que el mezclador reaccione de forma consecuente.

Por último se tratará de implementar un módulo de entrada y salida de vídeo por Network Device Interface -NDI-, un protocolo de red diseñado para la transmisión en vivo de señales multicámara. En esencia, se trata de un sustituto de la tradicional conexión entre cámara y CCU, ya que no solo soporta transmisión de vídeo sino que integra señales de retorno como tally o control remoto de cámaras robotizadas.

El vídeo que se transmite por NDI está comprimido. En concreto, usa fotogramas de tipo I y P -no B- de forma que la latencia de codificación-decodificación es mínima. Se dice que para una señal de 1080p30 se necesita un ancho de banda de 100Mbit/s, por lo que una única interfaz de Gigabit Ethernet -1000Mbit/s en dúplex- puede albergar varias señales de vídeo. Comparándola con una señal HD-SDI -capaz de transmitir la misma información en un ancho de banda de 1500Mbit/s- se obtiene un ratio de compresión de x15 -bastante ligero en comparación con otros codecs modernos-.

## Chapter 4

# Description

Lorem ipsum

## Chapter 5

# Results

Lorem ipsum

## Chapter 6

# Budget

Lorem ipsum

## Chapter 7

# Conclusions

Lorem ipsum