



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING or 371(c) DATE	GRP ART UNIT	FIL FEE REC'D	ATTY. DOCKET NO	TOT CLAIMS	IND CLAIMS
62/733,237	09/19/2018		280	46DG-286652		

CONFIRMATION NO. 2932

FILING RECEIPT



0000000102873437

137002

Sheppard Mullin Richter & Hampton LLP (GoPro)
650 Town Center Drive, 4th Floor
Costa Mesa, CA 92626

Date Mailed: 09/28/2018

Receipt is acknowledged of this provisional patent application. It will not be examined for patentability and will become abandoned not later than twelve months after its filing date. Any correspondence concerning the application must include the following identification information: the U.S. APPLICATION NUMBER, FILING DATE, NAME OF APPLICANT, and TITLE OF INVENTION. Fees transmitted by check or draft are subject to collection. Please verify the accuracy of the data presented on this receipt. **If an error is noted on this Filing Receipt, please submit a written request for a Filing Receipt Correction. Please provide a copy of this Filing Receipt with the changes noted thereon. If you received a "Notice to File Missing Parts" for this application, please submit any corrections to this Filing Receipt with your reply to the Notice. When the USPTO processes the reply to the Notice, the USPTO will generate another Filing Receipt incorporating the requested corrections**

Inventor(s)

César Douady, Les Molières, FRANCE;
Thomas Derbanne, Paris, FRANCE;
Maxim Karpushin, Paris, FRANCE;

Applicant(s)

GoPro, Inc., San Mateo, CA;

Power of Attorney:

D. Esplin--58297

Permission to Access Application via Priority Document Exchange: Yes

Permission to Access Search Results: Yes

Applicant may provide or rescind an authorization for access using Form PTO/SB/39 or Form PTO/SB/69 as appropriate.

If Required, Foreign Filing License Granted: 09/27/2018

The country code and number of your priority application, to be used for filing abroad under the Paris Convention, is **US 62/733,237**

Projected Publication Date: None, application is not eligible for pre-grant publication

Non-Publication Request: No

Early Publication Request: No

Title

SYSTEMS AND METHODS FOR STABILIZING VIDEOS

Statement under 37 CFR 1.55 or 1.78 for AIA (First Inventor to File) Transition Applications: No**PROTECTING YOUR INVENTION OUTSIDE THE UNITED STATES**

Since the rights granted by a U.S. patent extend only throughout the territory of the United States and have no effect in a foreign country, an inventor who wishes patent protection in another country must apply for a patent in a specific country or in regional patent offices. Applicants may wish to consider the filing of an international application under the Patent Cooperation Treaty (PCT). An international (PCT) application generally has the same effect as a regular national patent application in each PCT-member country. The PCT process **simplifies** the filing of patent applications on the same invention in member countries, but **does not result** in a grant of "an international patent" and does not eliminate the need of applicants to file additional documents and fees in countries where patent protection is desired.

Almost every country has its own patent law, and a person desiring a patent in a particular country must make an application for patent in that country in accordance with its particular laws. Since the laws of many countries differ in various respects from the patent law of the United States, applicants are advised to seek guidance from specific foreign countries to ensure that patent rights are not lost prematurely.

Applicants also are advised that in the case of inventions made in the United States, the Director of the USPTO must issue a license before applicants can apply for a patent in a foreign country. The filing of a U.S. patent application serves as a request for a foreign filing license. The application's filing receipt contains further information and guidance as to the status of applicant's license for foreign filing.

Applicants may wish to consult the USPTO booklet, "General Information Concerning Patents" (specifically, the section entitled "Treaties and Foreign Patents") for more information on timeframes and deadlines for filing foreign patent applications. The guide is available either by contacting the USPTO Contact Center at 800-786-9199, or it can be viewed on the USPTO website at <http://www.uspto.gov/web/offices/pac/doc/general/index.html>.

For information on preventing theft of your intellectual property (patents, trademarks and copyrights), you may wish to consult the U.S. Government website, <http://www.stopfakes.gov>. Part of a Department of Commerce initiative, this website includes self-help "toolkits" giving innovators guidance on how to protect intellectual property in specific countries such as China, Korea and Mexico. For questions regarding patent enforcement issues, applicants may call the U.S. Government hotline at 1-866-999-HALT (1-866-999-4258).

LICENSE FOR FOREIGN FILING UNDER**Title 35, United States Code, Section 184****Title 37, Code of Federal Regulations, 5.11 & 5.15****GRANTED**

The applicant has been granted a license under 35 U.S.C. 184, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" followed by a date appears on this form. Such licenses are issued in all applications where the conditions for issuance of a license have been met, regardless of whether or not a license may be required as set forth in 37 CFR 5.15. The scope and limitations of this license are set forth in 37 CFR 5.15(a) unless an earlier

license has been issued under 37 CFR 5.15(b). The license is subject to revocation upon written notification. The date indicated is the effective date of the license, unless an earlier license of similar scope has been granted under 37 CFR 5.13 or 5.14.

This license is to be retained by the licensee and may be used at any time on or after the effective date thereof unless it is revoked. This license is automatically transferred to any related applications(s) filed under 37 CFR 1.53(d). This license is not retroactive.

The grant of a license does not in any way lessen the responsibility of a licensee for the security of the subject matter as imposed by any Government contract or the provisions of existing laws relating to espionage and the national security or the export of technical data. Licensees should apprise themselves of current regulations especially with respect to certain countries, of other agencies, particularly the Office of Defense Trade Controls, Department of State (with respect to Arms, Munitions and Implements of War (22 CFR 121-128)); the Bureau of Industry and Security, Department of Commerce (15 CFR parts 730-774); the Office of Foreign Assets Control, Department of Treasury (31 CFR Parts 500+) and the Department of Energy.

NOT GRANTED

No license under 35 U.S.C. 184 has been granted at this time, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" DOES NOT appear on this form. Applicant may still petition for a license under 37 CFR 5.12, if a license is desired before the expiration of 6 months from the filing date of the application. If 6 months has lapsed from the filing date of this application and the licensee has not received any indication of a secrecy order under 35 U.S.C. 181, the licensee may foreign file the application pursuant to 37 CFR 5.15(b).

SelectUSA

The United States represents the largest, most dynamic marketplace in the world and is an unparalleled location for business investment, innovation, and commercialization of new technologies. The U.S. offers tremendous resources and advantages for those who invest and manufacture goods here. Through SelectUSA, our nation works to promote and facilitate business investment. SelectUSA provides information assistance to the international investor community; serves as an ombudsman for existing and potential investors; advocates on behalf of U.S. cities, states, and regions competing for global investment; and counsels U.S. economic development organizations on investment attraction best practices. To learn more about why the United States is the best country in the world to develop technology, manufacture products, deliver services, and grow your business, visit <http://www.SelectUSA.gov> or call +1-202-482-6800.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	46DG-286652
		Application Number	
Title of Invention	SYSTEMS AND METHODS FOR STABILIZING VIDEOS		
The application data sheet is part of the provisional or nonprovisional application for which it is being submitted. The following form contains the bibliographic data arranged in a format specified by the United States Patent and Trademark Office as outlined in 37 CFR 1.76. This document may be completed electronically and submitted to the Office in electronic format using the Electronic Filing System (EFS) or the document may be printed and included in a paper filed application.			

Secrecy Order 37 CFR 5.2:

☐ Portions or all of the application associated with this Application Data Sheet may fall under a Secrecy Order pursuant to 37 CFR 5.2 (Paper filers only. Applications that fall under Secrecy Order may not be filed electronically.)

Inventor Information:

Inventor 1				
Legal Name				
Prefix	Given Name	Middle Name	Family Name	Suffix
	César		Douady	
Residence Information (Select One) ☐ US Residency ☒ Non US Residency ☐ Active US Military Service				
City	Les Molières	Country of Residence ⁱ	FR	
Mailing Address of Inventor:				
Address 1		3000 Clearview Way		
Address 2				
City	San Mateo	State/Province	CA	
Postal Code	94402	Country ⁱ	US	
Inventor 2				
Legal Name				
Prefix	Given Name	Middle Name	Family Name	Suffix
	Thomas		Derbanne	
Residence Information (Select One) ☐ US Residency ☒ Non US Residency ☐ Active US Military Service				
City	Paris	Country of Residence ⁱ	FR	
Mailing Address of Inventor:				
Address 1		3000 Clearview Way		
Address 2				
City	San Mateo	State/Province	CA	
Postal Code	94402	Country ⁱ	US	
Inventor 3				
Legal Name				

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	46DG-286652
		Application Number	
Title of Invention	SYSTEMS AND METHODS FOR STABILIZING VIDEOS		

Prefix	Given Name	Middle Name	Family Name	Suffix
	Maxim		Karpushin	
Residence Information (Select One) ☐ US Residency ☒ Non US Residency ☐ Active US Military Service				
City	Paris	Country of Residence ⁱ	FR	
Mailing Address of Inventor:				
Address 1	3000 Clearview Way			
Address 2				
City	San Mateo	State/Province	CA	
Postal Code	94402	Country ⁱ	US	
All Inventors Must Be Listed - Additional Inventor Information blocks may be generated within this form by selecting the Add button. Add				

Correspondence Information:

Enter either Customer Number or complete the Correspondence Information section below. For further information see 37 CFR 1.33(a).			
☐ An Address is being provided for the correspondence information of this application.			
Customer Number	137002		
Email Address	dmipdocketing@sheppardmullin.com		Add Email Remove Email

Application Information:

Title of the Invention	SYSTEMS AND METHODS FOR STABILIZING VIDEOS		
Attorney Docket Number	46DG-286652	Small Entity Status Claimed	☐
Application Type	Provisional		
Subject Matter	Utility		
Total Number of Drawing Sheets (if any)		Suggested Figure for Publication (if any)	

Filing By Reference:

Only complete this section when filing an application by reference under 35 U.S.C. 111(c) and 37 CFR 1.57(a). Do not complete this section if application papers including a specification and any drawings are being filed. Any domestic benefit or foreign priority information must be provided in the appropriate section(s) below (i.e., "Domestic Benefit/National Stage Information" and "Foreign Priority Information").

For the purposes of a filing date under 37 CFR 1.53(b), the description and any drawings of the present application are replaced by this reference to the previously filed application, subject to conditions and requirements of 37 CFR 1.57(a).

Application number of the previously filed application	Filing date (YYYY-MM-DD)	Intellectual Property Authority or Country ⁱ

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	46DG-286652
		Application Number	
Title of Invention	SYSTEMS AND METHODS FOR STABILIZING VIDEOS		

Publication Information:

☐	Request Early Publication (Fee required at time of Request 37 CFR 1.219)
☒	Request Not to Publish. I hereby request that the attached application not be published under 35 U.S.C. 122(b) and certify that the invention disclosed in the attached application has not and will not be the subject of an application filed in another country, or under a multilateral international agreement, that requires publication at eighteen months after filing.

Representative Information:

Representative information should be provided for all practitioners having a power of attorney in the application. Providing this information in the Application Data Sheet does not constitute a power of attorney in the application (see 37 CFR 1.32). Either enter Customer Number or complete the Representative Name section below. If both sections are completed the customer Number will be used for the Representative Information during processing.

Please Select One:	☒ Customer Number	☐ US Patent Practitioner	☐ Limited Recognition (37 CFR 11.9)		
Customer Number	137002				
Prefix	Given Name	Middle Name	Family Name	Suffix	Remove
Registration Number					
Prefix	Given Name	Middle Name	Family Name	Suffix	Remove
Registration Number					
Additional Representative Information blocks may be generated within this form by selecting the Add button.					

Domestic Benefit/National Stage Information:

This section allows for the applicant to either claim benefit under 35 U.S.C. 119(e), 120, 121, 365(c), or 386(c) or indicate National Stage entry from a PCT application. Providing benefit claim information in the Application Data Sheet constitutes the specific reference required by 35 U.S.C. 119(e) or 120, and 37 CFR 1.78.

When referring to the current application, please leave the "Application Number" field blank.

Prior Application Status			Remove
Application Number	Continuity Type	Prior Application Number	Filing or 371(c) Date (YYYY-MM-DD)
Additional Domestic Benefit/National Stage Data may be generated within this form by selecting the Add button.			

Foreign Priority Information:

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	46DG-286652
		Application Number	
Title of Invention	SYSTEMS AND METHODS FOR STABILIZING VIDEOS		

This section allows for the applicant to claim priority to a foreign application. Providing this information in the application data sheet constitutes the claim for priority as required by 35 U.S.C. 119(b) and 37 CFR 1.55. When priority is claimed to a foreign application that is eligible for retrieval under the priority document exchange program (PDX) the information will be used by the Office to automatically attempt retrieval pursuant to 37 CFR 1.55(i)(1) and (2). Under the PDX program, applicant bears the ultimate responsibility for ensuring that a copy of the foreign application is received by the Office from the participating foreign intellectual property office, or a certified copy of the foreign priority application is filed, within the time period specified in 37 CFR 1.55(g)(1).

Application Number	Country ¹	Filing Date (YYYY-MM-DD)	Access Code ¹ (if applicable)

Additional Foreign Priority Data may be generated within this form by selecting the **Add** button.

Statement under 37 CFR 1.55 or 1.78 for AIA (First Inventor to File) Transition Applications

☐ This application (1) claims priority to or the benefit of an application filed before March 16, 2013 and (2) also contains, or contained at any time, a claim to a claimed invention that has an effective filing date on or after March 16, 2013.

NOTE: By providing this statement under 37 CFR 1.55 or 1.78, this application, with a filing date on or after March 16, 2013, will be examined under the first inventor to file provisions of the AIA.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	46DG-286652
		Application Number	
Title of Invention	SYSTEMS AND METHODS FOR STABILIZING VIDEOS		

Authorization or Opt-Out of Authorization to Permit Access:

When this Application Data Sheet is properly signed and filed with the application, applicant has provided written authority to permit a participating foreign intellectual property (IP) office access to the instant application-as-filed (see paragraph A in subsection 1 below) and the European Patent Office (EPO) access to any search results from the instant application (see paragraph B in subsection 1 below).

Should applicant choose not to provide an authorization identified in subsection 1 below, applicant **must opt-out** of the authorization by checking the corresponding box A or B or both in subsection 2 below.

NOTE: This section of the Application Data Sheet is **ONLY** reviewed and processed with the **INITIAL** filing of an application. After the initial filing of an application, an Application Data Sheet cannot be used to provide or rescind authorization for access by a foreign IP office(s). Instead, Form PTO/SB/39 or PTO/SB/69 must be used as appropriate.

1. Authorization to Permit Access by a Foreign Intellectual Property Office(s)

A. Priority Document Exchange (PDX) - Unless box A in subsection 2 (opt-out of authorization) is checked, the undersigned hereby **grants the USPTO authority** to provide the European Patent Office (EPO), the Japan Patent Office (JPO), the Korean Intellectual Property Office (KIPO), the State Intellectual Property Office of the People's Republic of China (SIPO), the World Intellectual Property Organization (WIPO), and any other foreign intellectual property office participating with the USPTO in a bilateral or multilateral priority document exchange agreement in which a foreign application claiming priority to the instant patent application is filed, access to: (1) the instant patent application-as-filed and its related bibliographic data, (2) any foreign or domestic application to which priority or benefit is claimed by the instant application and its related bibliographic data, and (3) the date of filing of this Authorization. See 37 CFR 1.14(h)(1).

B. Search Results from U.S. Application to EPO - Unless box B in subsection 2 (opt-out of authorization) is checked, the undersigned hereby **grants the USPTO authority** to provide the EPO access to the bibliographic data and search results from the instant patent application when a European patent application claiming priority to the instant patent application is filed. See 37 CFR 1.14(h)(2).

The applicant is reminded that the EPO's Rule 141(1) EPC (European Patent Convention) requires applicants to submit a copy of search results from the instant application without delay in a European patent application that claims priority to the instant application.

2. Opt-Out of Authorizations to Permit Access by a Foreign Intellectual Property Office(s)

☐ A. Applicant **DOES NOT** authorize the USPTO to permit a participating foreign IP office access to the instant application-as-filed. If this box is checked, the USPTO will not be providing a participating foreign IP office with any documents and information identified in subsection 1A above.

☐ B. Applicant **DOES NOT** authorize the USPTO to transmit to the EPO any search results from the instant patent application. If this box is checked, the USPTO will not be providing the EPO with search results from the instant application.

NOTE: Once the application has published or is otherwise publicly available, the USPTO may provide access to the application in accordance with 37 CFR 1.14.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	46DG-286652
		Application Number	
Title of Invention	SYSTEMS AND METHODS FOR STABILIZING VIDEOS		

Applicant Information:

Providing assignment information in this section does not substitute for compliance with any requirement of part 3 of Title 37 of CFR to have an assignment recorded by the Office.			
Applicant 1			
If the applicant is the inventor (or the remaining joint inventor or inventors under 37 CFR 1.45), this section should not be completed. The information to be provided in this section is the name and address of the legal representative who is the applicant under 37 CFR 1.43; or the name and address of the assignee, person to whom the inventor is under an obligation to assign the invention, or person who otherwise shows sufficient proprietary interest in the matter who is the applicant under 37 CFR 1.46. If the applicant is an applicant under 37 CFR 1.46 (assignee, person to whom the inventor is obligated to assign, or person who otherwise shows sufficient proprietary interest) together with one or more joint inventors, then the joint inventor or inventors who are also the applicant should be identified in this section.			
Clear			
☒ Assignee	☐ Legal Representative under 35 U.S.C. 117	☐ Joint Inventor	
☐ Person to whom the inventor is obligated to assign.	☐ Person who shows sufficient proprietary interest		
If applicant is the legal representative, indicate the authority to file the patent application, the inventor is:			
Name of the Deceased or Legally Incapacitated Inventor:			
If the Applicant is an Organization check here. ☒			
Organization Name	GoPro, Inc.		
Mailing Address Information For Applicant:			
Address 1	3000 Clearview Way		
Address 2			
City	San Mateo	State/Province	CA
Country ¹	US	Postal Code	94402
Phone Number		Fax Number	
Email Address			
Additional Applicant Data may be generated within this form by selecting the Add button.			

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	46DG-286652
		Application Number	
Title of Invention	SYSTEMS AND METHODS FOR STABILIZING VIDEOS		

Assignee Information including Non-Applicant Assignee Information:

Providing assignment information in this section does not substitute for compliance with any requirement of part 3 of Title 37 of CFR to have an assignment recorded by the Office.

Assignee 1				
Complete this section if assignee information, including non-applicant assignee information, is desired to be included on the patent application publication. An assignee-applicant identified in the "Applicant Information" section will appear on the patent application publication as an applicant. For an assignee-applicant, complete this section only if identification as an assignee is also desired on the patent application publication.				
If the Assignee or Non-Applicant Assignee is an Organization check here. ☐				
Prefix	Given Name	Middle Name	Family Name	Suffix
Mailing Address Information For Assignee including Non-Applicant Assignee:				
Address 1				
Address 2				
City		State/Province		
Country i		Postal Code		
Phone Number		Fax Number		
Email Address				
Additional Assignee or Non-Applicant Assignee Data may be generated within this form by selecting the Add button.				

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	46DG-286652
		Application Number	
Title of Invention	SYSTEMS AND METHODS FOR STABILIZING VIDEOS		

Signature:

NOTE: This Application Data Sheet must be signed in accordance with 37 CFR 1.33(b). **However, if this Application Data Sheet is submitted with the INITIAL filing of the application and either box A or B is not checked in subsection 2 of the "Authorization or Opt-Out of Authorization to Permit Access" section, then this form must also be signed in accordance with 37 CFR 1.14(c).**

This Application Data Sheet **must** be signed by a patent practitioner if one or more of the applicants is a **juristic entity** (e.g., corporation or association). If the applicant is two or more joint inventors, this form must be signed by a patent practitioner, **all** joint inventors who are the applicant, or one or more joint inventor-applicants who have been given power of attorney (e.g., see USPTO Form PTO/AIA/81) on behalf of **all** joint inventor-applicants.

See 37 CFR 1.4(d) for the manner of making signatures and certifications.

Signature	/D. Benjamin Esplin/			Date (YYYY-MM-DD)	
First Name	D. Benjamin	Last Name	Esplin	Registration Number	58297
Additional Signature may be generated within this form by selecting the Add button.					

Hypersmooth

- 1 Basic principles
 - 1.1 Definitions
 - 1.2 Hypersmooth philosophy
 - 1.3 Rotations, translations and full egomotion
 - 1.4 Rotations representation
 - 1.4.1 Notations
- 2 Problem position
 - 2.1 Objective function
 - 2.2 Constraints
- 3 Iterative minimization
 - 3.1 Update step
 - 3.1.1 First-order iterative minimization: gradient descent
 - 3.1.2 Second-order iterative minimization: Newton's method
 - 3.2 Projection step
 - 3.2.1 Lookup size–precision trade-off
- 4 Algorithm
 - 4.1 Algorithm I: Trajectory retrieval core
 - 4.1.1 Input
 - 4.1.2 Output
 - 4.1.3 Algorithm
 - 4.2 Algorithm II: Video stream stabilization
 - 4.2.1 Input
 - 4.2.2 Output
 - 4.2.3 Algorithm

Hypersmooth is a gyroscope-based video stabilization algorithm developed for hyperlapses.

This page tries to go beyond the algorithm description but also discusses key designing principles.

Basic principles

This section describes video stabilization philosophy and basic designing principles defining the application scope of Hypersmooth.

Definitions

A **camera model** is simply a mathematical model establishing correspondence between pixels and real world points. In the context of this article it is entirely described by the two projection instructions, from screen to world and vice versa, both taking into account not only intrinsic camera parameters but also the optical distortions. Since the 3D $\rightarrow$ 2D projection is an injection, i.e., whole ray from the optical center is projected onto a single screen point, we consider that the images of screen points are on the unit sphere S_1 surrounding the camera and centered at its optical center. $\mathbb{P}: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ $\mathbb{Q}: \mathbb{R}^2 \rightarrow S_1$ The two projecting instructions of a camera provide a complete description of its field of view. Precisely, we define the field of view as a set of points on S_1 , whose projections (images of $\mathbb{P}$) are visible, i.e., do not fall outside of the image plane Ω .

Source camera is the camera model corresponding to the input video. In what follows, the source camera might be rather *pinhole* or *hemispherical*. As an example, in case of Hero 5 capturing in wide FOV mode, the pinhole camera model applies with strong optical distortions, while a single sensor of Superbanc corresponds to a hemispherical camera model. Hypersmooth is thus designed to be applicable to both pinhole camera videos and punchouts from hemispheric input.

Output camera or just **crop** is the camera model corresponding to the output (stabilized) video. In the sense of the definitions given above, we can now formulate **the crop constraint**: for a given input video and camera models, **the field of view of the crop must remain a subset of the source camera field of view** at all times. In what follows, the crop is always a pinhole camera with any optical distortion model, totally independent of the source camera except that the crop constraint must be feasible.

Any quantity specifying how much the crop can move so that its field of view remains inside of the source field of view, is referred to as **stabilization margin**.

Hypersmooth philosophy

Camera motion is any change of its field of view. The source camera motion is measured by IMU and provided on the input. The crop is free to move; only the crop constraint is imposed. In such a setting, the goal of video stabilization is thus to find an optimal trajectory of the crop under that constraint, and reconstruct the images captured by the crop moving along this optimal trajectory, using the source camera video.

We distinguish two kinds of motion: intentional motion and the rest. Intentional motion is a meaningful phenomenon related the video semantics and expressing what the human being behind the camera actually wants to capture: a forward motion to follow an object of interest, panning to shoot a paysage, etc. Anything else is considered as noise caused by camera setup imperfections or its imposed motion: shaking hands when walking, vehicle vibration, etc.

Intentional motion is a hardly predictable low frequency phenomenon. Since it is all about low frequencies, a sufficiently long temporal horizon is needed to detect the intentional motion. In Hypersmooth, the optimal crop trajectory is searched for an entire time interval, and not just for the next sample as many other stabilization approaches seem to do. This introduces an unavoidable delay between the moment when a new motion sample is acquired and the optimal trajectory at that time moment is computed. The more this interval is, the visually better the approximated intentional motion is, however the more is the delay. The optimal trajectory might then be searched for a video stream in an overlapped-windowed fashion.

The intentional motion may not always be precisely reproduced by the crop, not only because of its mysterious nature. In the same way the noise might be too powerful with respect to the useful signal, the stabilization margin might not be sufficient to follow the intentional motion and do not violate the crop constraint. After all, the goal is not just to approach to the intentional motion as much as possible, but provide a visually pleasant crop trajectory. The human visual system seems prefer smooth low frequency motion (which fits well to the intentional motion nature). Based on this observation, we define the optimal trajectory simply as the one minimizing a certain score that expresses the amount of high frequencies. The key problem Hypersmooth tries to solve is thus a constrained minimization problem, whose variable is the trajectory itself and the objective function expresses that score. he minimization is preformed under the crop constraint.

Rotations, translations and full egomotion

The source camera motion might essentially be decomposed on translations of its optical center and rotations around it, giving 3 degrees of freedom each. Since the goal of video stabilization is not just discover the optimal crop trajectory but also render the crop images, it is important to take into account the reconstruction complexity corresponding to the kind of motion.

It is straightforward to see that, regardless the source camera projection model, if the two optical centers (source and crop) coincide, the crop image might be reconstructed by warping properly the source image captured at the same time. Let $\{I_{\text{source}}\}$ and $\{I_{\text{crop}}\}$ be the two images, $\{R\}$ a rotation matrix of the crop with respect to the source camera, and $\{(u,v) \in \Omega_{\text{crop}}\}$ a pixel location.

$$\{I_{\text{crop}}\}(u,v) = I_{\text{source}}(\text{left}(\text{mathbb{p}}_{\text{source}}\{R\}(u,v)\text{right}))$$

The crop constraint applies on $\{R\}$ and guarantees that the right side is defined for any meaningful pixel location $\{(u,v)\}$.

If there is also a translation, a similar formula does not apply. The pinhole projection model used on the crop side requires the relative scene depth at every point, which does not change with rotations but does change if a translation occurs. Moreover, a translation generally causes disocclusions, so that objects visible to the crop camera does not appear in the source camera image, and are therefore likely impossible to reconstruct from what has been captured by the source camera. The same applies in case of the full egomotion (rotation+translation).

Therefore, the stabilization transform in case of Hypersmooth is a rotation only, implying that the optical centers of the two cameras are at the same point. The searched optimal trajectory is thus the crop orientation with respect to the source camera.

Rotations representation

To deal with orientations and rotations numerically, we represent them in *axis-angle domain*:

- a rotation is represented by a vector co-directed with the rotation axis; the length of that vector is equal to the rotation angle in radians.
- A camera orientation is represented by a rotation to apply to the orientation origin.

This representation is compact, uniform and allows to approximate small rotations by additions/subtractions, so that finite difference approximations make sense. The latter is essential to deal with angular velocities and accelerations. The only issue of the axis-angle domain consists in its surjective nature with respect to $\{SO(3)\}$, so that a continuous set of orientations does not necessarily correspond to a continuous set of axis-angle vectors. However, a straightforward phase unwrapping procedure exists for the axis-angle representation, since the equivalence class for a given rotation representation $\{\vec{x}\}$ might be easily enumerated: $\{\tilde{\vec{x}}\} = \{\text{left}\{\vec{x}\} + 2\pi k \frac{\vec{x}}{\|\vec{x}\|}\}$: $k \in \mathbb{Z}$ } It is thus possible to find a close representation of a given rotation to a given reference orientation, which is naturally applies to temporal sequences of orientations we deal with.

Notations

We denote by $\{\vec{\cdot}\}$ a 3D-vector representing an orientation: $\{\vec{x}\} = (x, y, z)^T$. Bold literals specify N-dimensional vectors of a single orientation coordinate changing along the temporal dimension: $\{\mathbf{x}\} = (x_0, \dots, x_{N-1})^T$.

The mapping from axis-angle domain to rotation matrices (Rodrigues formula) is further referred to as $\{\rho\}$: $\{\rho(\vec{x})\} = R$.

Problem position

This section details the constrained optimization problem used to find the optimal trajectory.

Objective function

Let $\{\vec{x}_i\}_{i=0}^{N-1} = \{x_i, y_i, z_i\}_{i=0}^{N-1} \subset \mathbb{R}^3$ is an observed trajectory in axis-angle rotation representation. Every $\vec{x}_i$ represents a camera orientation when the center of the i -th frame has been captured. With positive indices we denote frames that are about to be stabilized (these and only these are the problem variables). We also consider orientations $\{\vec{x}_i\}_{i=-N_p}^{-1}$ corresponding to the frames in the fixed past, i.e. the ones that have already been stabilized and encoded. They are $(N_p = 2)$.

We minimize the following objective function: $J(\vec{x}_0, \dots, \vec{x}_{N-1}) = \alpha \sum_{i=-N_p}^{N-2} \|\vec{x}_{i+1} - \vec{x}_i\|^2 + \beta \sum_{i=-N_p+1}^{N-2} \|\vec{x}_{i+1} - 2\vec{x}_i + \vec{x}_{i-1}\|^2$. The objective function consists of two groups of terms weighted by two constants, (α) and (β) .

- From mechanical interpretation perspective, the first one expresses the (L_2) -norm of the interframe angular velocity, whereas the second one corresponds to the angular acceleration.
- From signal processing perspective, the two terms correspond to output power of two high-pass filters. This is coherent with our need to minimize high frequency motion.

The first observation consists in the fact that J is symmetrical and separable in the three coordinates of axis-angle domain: $J(\vec{x}_0, \dots, \vec{x}_{N-1}) = F_1(x_0, \dots, x_{N-1}) + F_1(y_0, \dots, y_{N-1}) + F_1(z_0, \dots, z_{N-1})$. Here F_1 is of a similar form: we split fixed past terms accordingly. $F_1(x_0, \dots, x_{N-1}) = \alpha \sum_{i=-N_p}^{N-2} (x_{i+1} - x_i)^2 + \beta \sum_{i=-N_p+1}^{N-2} (x_{i+1} - 2x_i + x_{i-1})^2$. Further we investigate properties of F_1 instead the complete objective function J .

From algebraic point of view, F_1 is a second order polynomial, and may thus be rewritten in the following form $F_1(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{H} \mathbf{x} + \mathbf{w}^T \mathbf{x} + c$. Here (c) is a scalar and $(\mathbf{w})$ is a vector, both depending on fixed past. $(\mathbf{H})$ has a quasi-Toeplitz form as follows. $\mathbf{H} = \begin{pmatrix} 4\alpha + 12\beta & -2\alpha - 8\beta & 2\beta & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4\alpha + 12\beta & -2\alpha - 8\beta & 2\beta & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4\alpha + 12\beta & -2\alpha - 8\beta & 2\beta & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4\alpha + 12\beta & -2\alpha - 8\beta & 2\beta & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4\alpha + 12\beta & -2\alpha - 8\beta & 2\beta & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4\alpha + 12\beta & -2\alpha - 8\beta & 2\beta & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4\alpha + 12\beta & -2\alpha - 8\beta & 2\beta & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4\alpha + 12\beta & -2\alpha - 8\beta & 2\beta & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4\alpha + 12\beta & -2\alpha - 8\beta & 2\beta & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4\alpha + 12\beta & -2\alpha - 8\beta & 2\beta & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4\alpha + 12\beta & -2\alpha - 8\beta & 2\beta & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4\alpha + 12\beta & -2\alpha - 8\beta & 2\beta & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4\alpha + 12\beta & -2\alpha - 8\beta & 2\beta & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4\alpha + 12\beta & -2\alpha - 8\beta & 2\beta & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4\alpha + 12\beta & -2\alpha - 8\beta & 2\beta & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4\alpha + 12\beta & -2\alpha - 8\beta & 2\beta & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4\alpha + 12\beta & -2\alpha - 8\beta & 2\beta & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4\alpha + 12\beta & -2\alpha - 8\beta & 2\beta \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4\alpha + 12\beta & -2\alpha - 8\beta \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4\alpha + 12\beta \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$. The matrix notation allows to express the gradient of F_1 in a compact way: $\nabla F_1(\mathbf{x}) = \mathbf{H} \mathbf{x} + \mathbf{w}$. It is therefore straightforward to end up with the Hessian matrix of F_1 equal to $(\mathbf{H})$.

We fix $(\alpha = \beta = 1)$ for the moment to investigate numeric properties of the Hessian matrix $(\mathbf{H})$. The figure below illustrates some statistics on its eigenvalues.

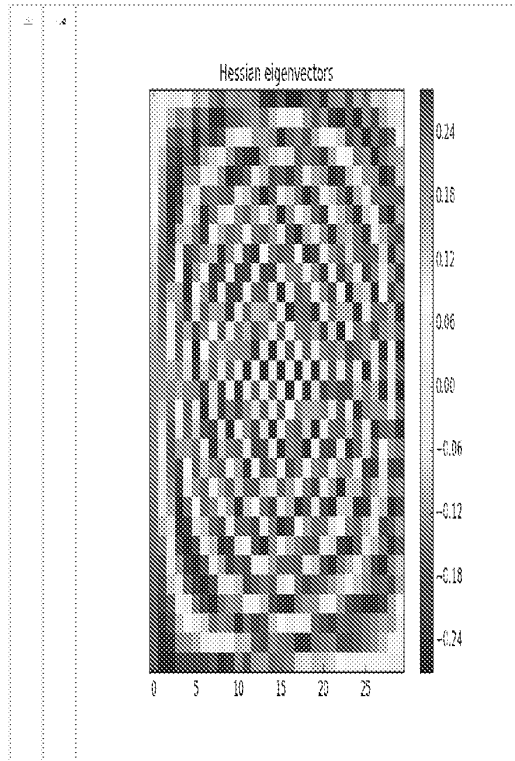


Figure: Hessian condition number for different input length N (left), Hessian spectrum for input of $N=30$ (middle) and corresponding eigenvectors (right).

Three key properties of the Hessian matrix matter for its future use for the stabilization.

1. Thanks to the objective function simplicity, the Hessian depends on α and β , but **does not depend on data** ($\mathbf{x}$). This allows to precompute any its characteristics offline.
2. Hessian is positive definite, as all its eigenvalues are positive. This implies the following.
 1. F_1 is convex.
 2. Equation $\nabla F_1(\mathbf{x}) = H \mathbf{x} + \mathbf{w} = 0$ always has a unique solution: there is only one critical point.
 3. F_1 reaches its global minimum at that point.
3. In the tested spectrum of N , the maximum eigenvalue increases monotonically, and remains quite close but does not go beyond 40 ($\lambda_{\max}=39.25$) for $N=15$, $\lambda_{\max}=39.93$ for $N=50$).

Constraints

The only constraint of the optimization problem we are about to pose corresponds to the crop constraint. To formulate it, we first sample the crop screen perimeter, obtaining a set of points $B \subset S_1$. We took four crop screen corners and a point in the middle of every screen edge, which gives another four points. An epsilon distance outwards was added to the latter four due to the boundary curvature specificity, to guarantee the crop constraint respected. This set is fixed forever and does depend only on the intrinsic parameters of the crop and the two distortion models. The constraint on a given rotation R is then defined as follows: $\mathcal{T}(R) = \begin{cases} 1, & \mathbf{p}_{\text{source}} \in R \mathbf{q}_{\text{crop}}(u,v) \in \Omega_{\text{source}} \\ 0, & \text{for all } (u,v) \in B \setminus \Omega_{\text{source}} \end{cases}$

Since $\mathbf{x}_i$ are camera orientations in absolute coordinates, we need to relativise the stabilization transform in order to check whether the constraint applies. We store the input trajectory $\left\{ \mathbf{x}_i \right\}_{i=0}^{N-1} = \left\{ \mathbf{x}_i \right\}_{i=0}^{N-1}$ and apply the constraints test $\mathcal{T}$ as follows: $\mathcal{T}(\left\{ \mathbf{x}_i \right\}) = \mathcal{T}(\left\{ \mathbf{x}_i \right\})$

The feasible set is then engendered by the indicator $\mathcal{X}_i = \left\{ \mathbf{x}_i : \mathcal{T}(\left\{ \mathbf{x}_i \right\}) = 1 \right\}$. The entire problem feasible set is given by Cartesian product of $\mathcal{X}_i$ over all $i = 0, \dots, N-1$.

There are few things we can say about $\mathcal{X}$.

1. $\mathcal{X}$ depends not only on the camera models, but also on the source camera orientation. However, it is independent of the fixed past.
2. The constraint is applied in a pointwise fashion, i.e., every trajectory point (x_i, y_i, z_i) is constrained separately from the others. This is due to the Cartesian product in the definition.
 - This is somewhat dual to the objective function separability, which is separable in "orthogonal direction", i.e., along $\mathbf{i}$.

The same feasible set may be defined differently through a continuous function taking zero value inside $\mathcal{X}$ and progressively increasing outside. For example, we may measure the distance from every crop screen corner to the source screen perimeter and shrink it such that it remains zero if the corner falls inside the source FOV. While it is possible to define a differentiable penalization score like this, the analytic expression of its derivative is horribly complex: an unsimplified symbolic expression of one coordinate derivative for hemispherical input gives 568 Kb of code in LaTeX notation; its symbolic simplification does not end. This makes unfeasible to integrate the constraints through Lagrange multipliers.

The optimization problem of finding the optimal crop trajectory can now be fully formulated.
$$F(\vec{x}_0, \dots, \vec{x}_{N-1}) = \alpha \sum_{i=-N_p}^{N-2} \left((x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2 + (z_{i+1} - z_i)^2 \right) + \beta \sum_{i=-N_p+1}^{N-2} \left((x_{i+1} - 2x_i + x_{i-1})^2 + (y_{i+1} - 2y_i + y_{i-1})^2 + (z_{i+1} - 2z_i + z_{i-1})^2 \right) \rightarrow \min \{ (x_i, y_i, z_i) = \vec{x}_i \in \mathcal{X}_i, i = 0, \dots, N-1 \}$$

Iterative minimization

The preceding discussion may be concluded as follows: while the objective function is rather simple, the constraints are rather complicated. In case of global optimization, an iterative minimization of such a simple objective function does not make sense, but in the constrained case we end up with an iterative approach due to the constraints complexity.

A classic approach would consist in iterating the following two steps altering the current estimate.

- The current estimate is updated regardless the constraints in a way that the objective function value decreases. This is the **update step**.
- The constraints are applied: the current estimate is tested whether it remains inside the feasible set. If is not the case, it is reprojected back onto it. This is **projection step**.

The both steps assume a small change of the current estimate so that it converges slowly to the optimal solution.

Since the problem variables express explicitly the trajectory, at any moment the current estimate is a crop trajectory which is both feasible and smooth (smoother than the input one). This is advantageous for our time-constrained usecase: we may do as many iterations as possible and stop at any moment, getting a smoothed crop trajectory. An early stopping criteria might be considered, e.g. if the estimate change on the update step becomes too small or zero.

In what follows we describe the both steps as proposed in Hypersmooth.

Update step

The following discussion exploits the objective function separability, i.e., we deal with F_1 along each orientation dimension $\langle x \rangle$, $\langle y \rangle$ and $\langle z \rangle$.

First-order iterative minimization: gradient descent

Classic gradient would consist in iterating the following formula: $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \gamma \nabla F_1(\mathbf{x}^{(k)})$. γ is a constant step. The iterations converge if and only if $\gamma < \frac{1}{\lambda_{\max}(H)}$.

The advantage of this simple gradient descent is in the simplicity of the update step. It is lightweight, easily implementable and vectorizable. The major limitation is the convergence speed: this is a "first order" approach meaning that the objective function error on iteration $\langle k \rangle$ is bounded by $O(\frac{1}{k})$. The convergence rate strongly depends on the Hessian conditionnedness: the greater is its eigenvalues ratio, the less optimal is the update step direction given by the gradient. A lot of iterations may be needed to get close to the minimum. A number of techniques exists to speedup the gradient descent, the ones we use are detailed further.

Second-order iterative minimization: Newton's method

To speedup the convergence rate, a second order iterative minimization is of interest. The multivariate generalization of the Newton's method consists in iterating the following formula: $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - H^{-1} \nabla F_1(\mathbf{x}^{(k)})$. While the classic gradient descent approximates the objective function by a plane, the Newton's method assumes a quadratic surface. This allows to choose the update step direction better, eliminating the impact of the Hessian conditionnedness and guaranteeing the objective function error bounded by $O(\frac{1}{k^2})$.

Projection step

As described on that page, several approaches to compute the projection step have been explored.

- The feasible set may be approximated in a convenient way to compute the orthogonal projection. The projection itself is then precisely computed on the approximated feasible set. In the projected gradient descent we fit a polytope to the feasible set. The major limitation of this technique is that it remains too slow, even after the polytope is tweaked in a special way and the projection is computed in $\mathcal{O}(\log M)$ operations for M -vertex polytope.
- The projection may be approximated keeping the feasible set as. In the bounded gradient descent we are directly invoking $\text{proj}_{\mathcal{X}}(\vec{x}_i)$ to approximate the projection. Counterintuitively, this turns out to be faster. However, the way the projection is approximated is too limiting, so that the achievable stabilization margin is shrinked with respect to the potential one.

The eventually proposed approach approximates both the feasible set and the projection, but in a more efficient way, becoming both faster and more accurate in terms of the use of stabilization margin than the two other techniques. This is based on the observation that there is a meaningful subset of $\mathcal{X}_i$ is radially convex (or a "star-shaped" if you like the math to be fancy): for an arbitrary rotation axis there should be a maximum rotation angle, such that for any rotation above this angle the crop field of view remains inside the source crop. An exception is the unbounded rolling in case of a hemispherical source camera. Another degenerative example is a square crop in a squared-FOV source camera, which could not be fitted if rotated by 45° , but could be if rotated 90° or 270° . Meaningful rotations are the ones that allow to render the trajectory smooth; such degenerative cases are thus out of interest.

This maximum angle could be easily sampled for a given direction $\vec{x}$, $\|\vec{x}\| = 1$, evaluating $\langle \eta \vec{x}, \vec{x} \rangle$ in a dichotomic fashion with respect to a positive scalar η . The second key observation consists in the fact that this angle, as any other function defined on $\mathcal{S}_1$, is easily lookupable through a cubemap-like table: we reproject the axis on the unit cube and hit the table corresponding to the face the axis crosses. We proceed as follows.

The lookup is computed offline and depends only on the two camera models.

- We sample limiting rotation angles for few directions falling in a given cell of the lookup table.
- We fit a plane to the sampled point cloud and adjust its constant coefficient in such a way that all the rotations are in the outer halfspace with respect to that plane (the inner one would be the one including the origin). This approximates the tangent plane to the feasible set nearby the cell center. By convention, the plane normal is oriented towards that outer halfspace.
- The four plane coefficients are stored in the lookup then.

Online projection step consists then in the following procedure, applied consecutively to every $\vec{x}_i$, $i = 0, \dots, N - 1$.

- We first compute $\vec{\delta}_i = \rho^{-1} \left(\rho \vec{x}_i \right) \left[\rho \left(\vec{x}_i^{(0)} \right) \right] \text{right}^T \text{right} \equiv \vec{x}_i \circ \left(-\vec{x}_i^{(0)} \right)$. A straightforward computation of this quantity (conversion to rotation matrices and backward) is heavy. According to our tests, finely optimized formula derived through symbolic computation from the Rodrigues formula is about 10.5 times faster than the direct computation (double precision floating point, 64 bit machine, OpenCV for vector arithmetic and Rodrigues transform baseline). Another option is to pass by unit quaternions, since conversion between axis-angle vectors and unit quaternions does not require too much computing, as well as the quaternions composition. Intuitively, our computation scheme obtained with symbolic computing seems close to one would get through the quaternions.
- Then we hit the lookup with $\vec{\delta}_i$, i.e., measure the signed distance from $\vec{\delta}_i$ to the corresponding plane.
- If the distance is negative, the rotation angle for the given rotation (equal to $\|\vec{\delta}_i\|$ due to the use of the axis-angle representation) does not exceed the limiting one, so the rotation is feasible.
- Otherwise, a reprojection on the feasible set is required. $\vec{\delta}_i$ is first projected on the tangent plane. Its projection would typically fall out from the feasible set, in which case we hit the lookup one more time searching for the intersection of the reprojected rotation and corresponding plane. The resulting point is after all taken as the projection approximation and used to update $\vec{x}_i$, bringing it back to the feasible set (likely not too far from the actual orthogonal projection). Here, the axis-angle vector composition operation is applied again: $\vec{x}_i = \vec{\delta}_i \circ \vec{x}_i^{(0)}$

Lookup size–precision trade-off

The cubemap lookup allows thus to test efficiently whether a rotation is feasible. However, it approximates the feasible set boundary by planes that are shrunk in a way that all the points used for fitting remain beyond. This this causes a small loss of the stabilization margin.

The loss is controlled per cell and is measured as the maximum distance over all the points used to fit a plane in a given cell, to the resulting boundary plane. Corresponding statistics are reported below. As it could be seen in the Table and the Figure below, the maximum loss is marginal even for a relatively small lookup (10*10 cells per cube map), and its mean value is shifted towards zero.

Resolution per face	Total lookup size, bytes	Max loss per cell, °	Avg. loss per cell, °	Loss std, °
10*10	9 600	0.73	0.0268	0.0867
20*20	38 400	0.49	0.0083	0.0391
50*50	240 000	0.19	0.0013	0.0095
100*100	960 000	0.10	0.0004	0.0033

Table: Lookup memory footprint and stabilization margin loss for different cubemap resolutions (number of cells per face). Singular precision floating point arithmetic is assumed to store the plane coefficient (32 bits per plane coefficient).

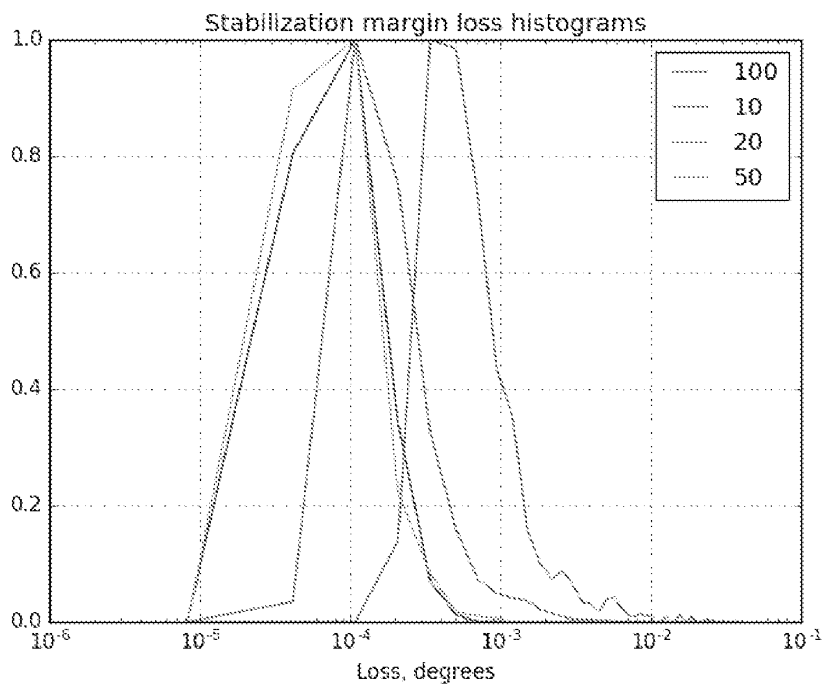


Figure: Normalized histograms of the maximum per-cell stabilization precision loss for different lookup resolutions.

Algorithm

Algorithm I: Trajectory retrieval core

Input

1. Data

1. Input trajectory: source camera orientation at every frame
 - Axis-angle representation used
 - The input is assumed phase-unwrapped
 - The input must have at most $\lfloor N \rfloor$ samples
1. Bootstrap trajectory (pass the input trajectory if nothing else is available)
2. Fixed past (a source camera trajectory that the input trajectory is a continuation of, empty if not available)
2. Control parameters
 1. Number of iterations $\lfloor N_{it,1} \rfloor$, $\lfloor N_{it,2} \rfloor$, $\lfloor N_{it,3} \rfloor$
 - Recommended: $\lfloor N_{it,1} \rfloor = \lfloor N_{it,3} \rfloor = \frac{N_{it}}{3}$, $\lfloor N_{it,2} \rfloor = \frac{2 N_{it}}{3}$ for $\lfloor N_{it} \rfloor = 100$

Output

1. Stabilized crop trajectory: a camera orientation sample per frame as in the input trajectory
 - Axis-angle representation is used
 - The output trajectory is phase-unwrapped

Algorithm

1. Prepare projection: copy the input trajectory into a projection buffer
2. Bootstrap: take the bootstrap trajectory and pad it if it is shorter than $\lfloor N \rfloor$. This will be the current guess.
 1. If there is no fixed past, extrapolate on the left using the first sample
 2. Otherwise extrapolate on the right using the last sample
3. Compute constant step terms for both fine and coarse scales
4. Fine scale: repeat $\lfloor N_{it,1} \rfloor$ times
 1. For every point in current guess, separately and consecutively (in the increasing time order)
 1. Compute gradient using a precomputed power of Hessian and the constant gradient term
 2. Compute the update: add the gradient term to the current guess
 3. Compute the stabilizing rotation $\lfloor \delta \rfloor$ through axis-angle composition ($\circ$ operation)
 4. Hit the boundary lookup with $\lfloor \delta \rfloor$: get the limiting plane
 5. Compute the distance from $\lfloor \delta \rfloor$ to the limiting plane
 6. If it is negative, assign the update to the current guess (at the current point)
 7. Otherwise,
 1. project $\lfloor \delta \rfloor$ onto the limiting plane,
 2. hit the boundary lookup for the corresponding limiting plane (*crashsite*),
 3. scale the projection of $\lfloor \delta \rfloor$ so that it falls on the crashsite,
 4. compute new update value through axis-angle composition (compose the input position with the obtained projection),
 5. assign the result to the current guess (at the current point)
5. If $\lfloor N_{it,2} \rfloor > 0$, subsample the current guess
6. Coarse scale: repeat $\lfloor N_{it,2} \rfloor$ times
 1. Repeat 4.a using the coarse scale Hessian and constant gradient step, on subsampled current guess
7. If $\lfloor N_{it,2} \rfloor > 0$, upsample the subsampled current guess (linear interpolation used) and assign it to the current guess
8. Fine scale again: repeat $\lfloor N_{it,3} \rfloor$ times
 1. Repeat 4.a as is
9. Unpad the current guess and write it to the output

Algorithm II: Video stream stabilization

Input

1. Video stream: frames and their timestamps becoming available in time
 - "Frame timestamp" is the moment when the center scanline is captured
2. Gyroscope stream: packets of samples becoming available in time
 - Gyroscope is assumed synchronized with video stream
3. Control parameters

Output

1. Stabilized video
2. Stabilized crop orientation in time (if needed)

Algorithm

The offline phase consists in initializing all the stuff used on the online phase.

The online phase is invoked on every new video frame.

1. Integrate gyro samples to reach the frame timestamp; this can be invoked separately
2. Put the frame into the circular buffer; the source camera orientation obtained from the integrated gyro stream and the timestamp always go together
3. Increase $\backslash(n_phi\backslash)$ (initialized to zero at the beginning)
4. If there are $\backslash(N\backslash)$ frames in the buffer and $\backslash(n_phi \geq \backslashPhi\backslash)$, go
 1. Unwrap the input trajectory
 2. Prepare the fixed past: take the crop orientation history and unwrap it backwards with respect to the first sample of the input trajectory; take at most $\backslash(N_p\backslash)$ samples
 - If there is no history yet, consider the fixed past empty
 3. Prepare the bootstrap: take the crop orientation history (as many samples as possible), fill the rest with the input samples according to their timestamps and unwrap
 - If there is no history yet, consider the bootstrap equal to the input trajectory
 4. Invoke Algorithm I
 5. Store the output stabilized trajectory to the orientation history
 6. Set $\backslash(n_phi\backslash)$ to zero
 7. Warp and encode frames in the head of the buffer, at least 1, at most $\backslash(n_phi\backslash)$; for every such frame
 1. Compute the rotation corresponding to pass from the source orientation to the stabilized one
 2. Sample the crop field of view: take a grid of points
 3. Rotate the grid accordingly (apply the rotation to every its point)
 4. Invoke ERS correction, if one wants to
 5. Project the grid on the source camera image plane
 6. Invoke the warp, passing the grid and the corresponding input image
 7. Send the warped image to the encoder

After the video stream is over, repeat 4.a – 4.f once, and 4.g.i – 4.g.vii for all the remaining frames in the buffer.

Hypersmooth tuning entries overview

Hypersmooth is an EIS algorithm destined for Tavarua. It consists of (1) a main deferred stabilizer providing a highly attenuated motion and (2) a secondary causal lightweight stabilizer used to stabilize the display image for life view.

This page explains the available tuning parameters of the both stabilizers, their applicable ranges and their impact on the stabilization behavior.

 The parameter values may be set using t-commands. Setting a parameter value out of its corresponding valid range would cause a failure to configure the stabilizers. The EIS functionality would not work then.

Introduction

Within this page we use the following terminology.

Stabilizers deal with motion. The *input motion* is measured by the IMU of the camera (the gyroscope, precisely). Stabilized image may be seen as an image captured by a virtual camera (*crop*), whose FOV is included into the FOV of the real camera (the inclusion is defined in the sense of the set theory, while the FOV itself is a set of points on a unit sphere visible by the camera). The FOV inclusion expresses *crop constraint*.

The goal for a stabilizer becomes therefore to compute a *crop trajectory* that respects the crop constraint and provides a visually smooth sequence of images.

Main stabilizer tuning parameters

Main stabilization runs in deferred mode, providing a highly attenuated motion under the crop constraint. In the way it is currently integrated on Tavarua, its output is encoded in MP4 and LRV files and is available through the HDMI connection. The approach behind is based on an optimization problem whose objective function is related to physical quantities expressing the amount of global motion in the video.

(1) Weight balance

Weight balance is the main parameter impacting the stabilization style, i.e., defining which motion should actually be considered as optimal. It is a scalar parameter in $[0,1]$ range, internally quantized to 5 uniformly spaced values: 0, 0.25, 0.5, 0.75, 1.

- **When set to "0"**, the angular velocity is solely minimized in the resulting motion.
This would tend to piecewise-linear sharp-cornered trajectories and prefer to keep a steady position when possible.
 It actually minimizes the apparent motion between images.
 Due to its piecewise-linear nature, the resulting motion may appear unnatural, since in a Newtonian world the velocity is everywhere differentiable quantity.
- **When set to "1"**, the angular acceleration is solely minimized in the resulting motion.
 This would result in smooth, fully differentiable Newtonian motion with no corner points.
 Due to the fact that minimizing acceleration corresponds to keeping a constant velocity (and not necessarily zero), this setting generally produces more motion.
 Depending on the other parameters, acceleration-only minimization may render low frequency oscillations when standing still after a turn. This setting is thus not recommended for use in production.
- When set in-between, both minimization components are weighted accordingly, making "round corners" while reducing the apparent motion.

Default value: 0.25 for MP, updated to 0.5 for Launch.

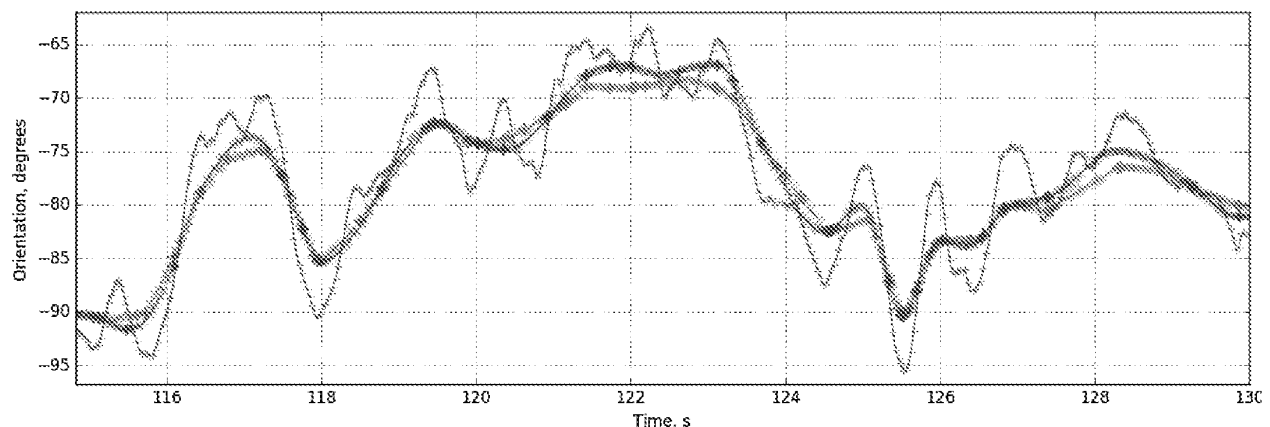


Figure 1. Weight balance impact on the stabilization behavior. Green curve is the input motion (walking, handheld camera, orientation around vertical axis). Stabilized trajectories with 0% and 100% weight balance are in red and blue respectively. Crosses on stabilized trajectories denote constraint boundary hitting instants.

(2) Low-light high-pass (motion blur masking control)

Taking into account the exposure time of every frame, Hypersmooth decomposes the observable camera motion in two categories.

- **Inter-frame motion** is the motion between frame exposure periods. This motion can be attenuated completely up to the cropping constraint.
- **Intra-frame motion** is the motion underwent by the camera during the exposure period of every frame. This motion is already encrusted in the image through the motion blur, and its attenuation causes motion blur pulsations producing a bad visual impression.

Arguably the only way to deal with the latter is to render the trajectory representing the same motion during the exposure periods, i.e., in a way that the camera always moves in the same direction and with the same speed during the exposure phase of every frame, and might then stand still between consecutive frames.

There is a positive scalar parameter controlling the gain of the corresponding high-pass filter.

- When set to 0, no motion blur compensation is done; the observed motion is fully attenuated and no camera motion during the exposure periods is taken into account. The output motion is fully smooth. This would end with a similar result as early Chopes.
- When set to 1, the entire camera displacement during the exposure periods is preserved. The output motion contains high frequencies that mask motion blur.

This parameter does not need to evolve in function of the exposure time.

- When a highly illuminated scene is being shot, the exposure time is short enough to render the intra-frame displacement negligible.
- When the camera does not move much, the intra-frame displacement is limited regardless the exposure time.

Default value: 1.0 for MP, updated to 0.8 for Launch.

(3) Stickiness

High frequencies are normally completely filtered out by Hypersmooth, except if the parameter above allows a high-pass. Similarly, a low-pass filter might be enabled to keep a low frequency component present in the original motion. Stickiness is a scalar parameter in $[0, 1]$ range controlling this behavior.

 For a regular shaky motion this parameter has no impact. It makes difference only for low-frequency input motion usecases, such as a camera mounted on a UAV, where the stabilized point of view may remain the same for a long time.

- **Full stickiness (1)** renders the stabilizer "lazy" so that it keeps its attitude depending only on the past signal, unless the observed future samples exhibit a clear need to change it, i.e., the crop will fall out of the source FOV. The "attitude" depends mainly on the weight balance discussed above.
 - In case of zero weight balance, the stabilizer slows down and keeps the same position as much as possible.
 - In case of the maximum weight balance (1), it tends to a constant velocity.

 This extreme setting renders the stabilizer sensitive to the gyroscope bias (see Figure 2). It is thus not recommended for use in production.

- **Zero stickiness** would enable a low-pass filter in a way that the stabilizer behavior does not depend only on the past samples, but always impacted by the future ones. The exact behavior of the filter (e.g., frequency response) is not specified and depends on the other parameters.

Default value: 0.5.

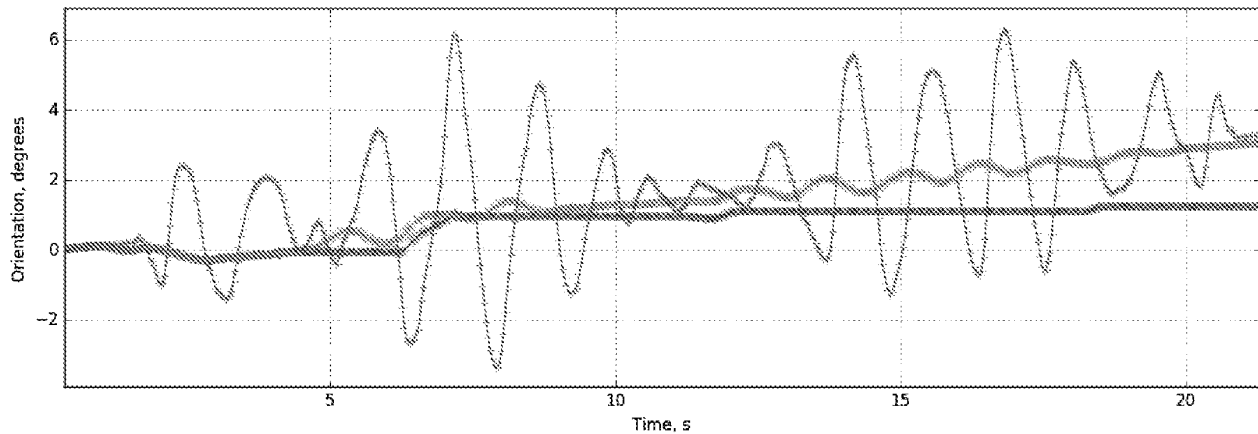


Figure 2. Stickiness impact on the stabilization behavior. Green curve is the input motion (UAV, orientation around vertical axis). Stabilized trajectory with 0% and 100% stickiness are in red and blue respectively.

Display stabilizer tuning parameters

Due to its deferred nature, main stabilizer could not be used to stabilize the display image without latency. In order to provide the stabilized live view, Hypersmooth includes a second lightweight causal (using only past samples) stabilizer. Its algorithm is based on a different principle, rather heuristic, and does not intend to render an optimal stabilized motion. Therefore it naturally exhibits more tuning entries.

The key principle the display stabilization is based on consists in estimating whether the currently observed camera motion is intentional or not. Intentional motion is then preserved, while unintentional motion is attenuated. This is handled through *stabilizer state variable*, which expresses the degree of intentionality in time. The state estimation is based mainly on the angular acceleration.

- When the input angular acceleration increases too much and goes beyond a threshold, the state becomes "fully unintentional". No attenuation applies in this state.
- When the input angular acceleration becomes too low, the state becomes "fully intentional". Full attenuation applies in this state.
- Otherwise, the state is interpolated in-between.
 - When the acceleration increases, the state updates instantly: the attenuation level increases to a target interpolated level immediately.
 - When the acceleration drops, the state updates progressively: the attenuation level decreases slowly.

An example is given on Fig. 3.

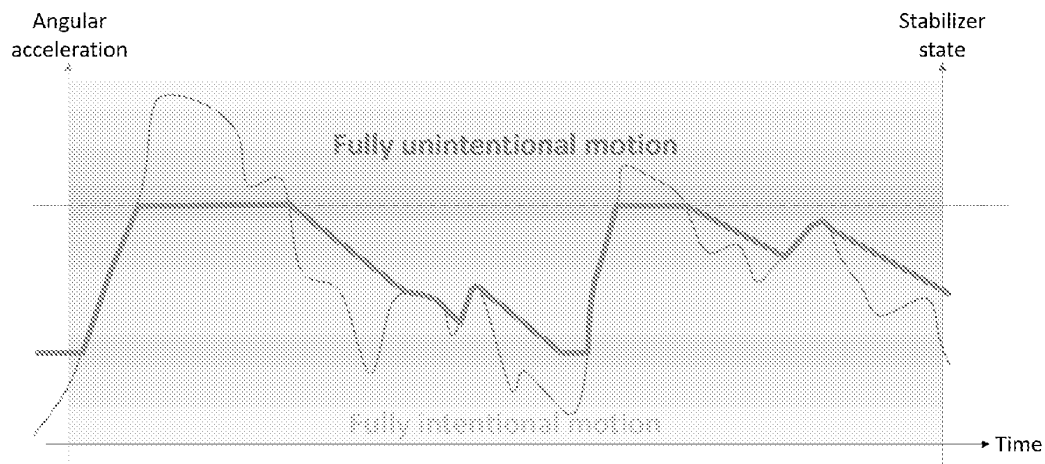


Figure 3. Display stabilizer state evolution (in blue) in function of observed angular acceleration (in orange).

State control parameters

Acceleration thresholds

Two scalar parameters specify minimum and maximum acceleration thresholds fully intentional and fully unintentional motion.

Default values: 200°/s² and 2000°/s² respectively.

 When configuring the stabilizer, the passed intentional (minimum) acceleration threshold should always be less than the unintentional (maximum) one. An error occurs otherwise.

Centering force

Crop motion follows the input through a recentering mechanism: the angular distance from the crop towards the FOV center is multiplied by a centering force factor to drop to zero progressively and thus recenter the crop.

The stabilizer state is mapped to the centering force in a linear fashion.

- In the fully unintentional state the centering force is null: the crop would try to keep its position as much as possible and will only move when the crop constraint is not respected.
- In the fully intentional state, the centering force takes its maximum value, which is a tuning parameter in [0, 1] range: the crop rotation is reduced by that value for every frame, exponentially decreasing to zero.

Default value: 0.1.

Inertia control parameters

An inertia component is added to the crop speed to render its motion more natural. It consists in keeping a portion of the current speed when changing computing its next position. This is controlled by two scalar parameters in [0, 0.5] range, corresponding to the fully intentional and fully unintentional motion. Intermediate inertia values are interpolated in function of the stabilizer state.

Default values: 0.25 and 0.49 respectively.

Anticipation limit

Limits apply to reduce the maximum angular distance from the crop orientation towards the FOV center. These limits may be used to render the stabilization more reactive in modes exposing a lot of stabilization margin, such as linear FOV and zoomed modes.

- Slant anticipation limit specifies the maximum deviation in degrees in horizontal and vertical directions (maximum panning and tilting respectively).
- Roll anticipation limit specifies the maximum rotation around the optical axis.

Default values: 10° and 5° respectively.

T-commands

The following commands are implemented to allow the stabilizer tuning. They follow the order they are described above. The passed values match the default ones recommended for Launch.

 These t-commands might be put in autoexec.sh. They take effect, once the capture mode is changed or the recording starts or stops.

```
t frw imgcap eis hs_tuning wbal 0.50
t frw imgcap eis hs_tuning sticky 0.50
t frw imgcap eis hs_tuning blur 0.80
t frw imgcap eis hs_tuning minacc 200
t frw imgcap eis hs_tuning maxacc 2000
t frw imgcap eis hs_tuning cforce 0.100
t frw imgcap eis hs_tuning iinert 0.250
t frw imgcap eis hs_tuning uinert 0.490
t frw imgcap eis hs_tuning slant 10.00
t frw imgcap eis hs_tuning roll 5.00
```

Appendix A

SYSTEMS AND METHODS FOR STABILIZING VIDEOS***FIELD***

(01) This disclosure relates to stabilizing videos using positions of an image capture device during a capture duration.

BACKGROUND

(02) A video may have been captured by an image capture device in motion. The motion of the image capture device during the capture of the video may cause the video to appear jerky/shaky.

SUMMARY

(03) This disclosure relates to stabilizing videos. Images with an optical field of view may be captured by an image capture device during a capture duration. Image information defining the images, position information characterizing positions of the image capture device at different moments within the capture duration, and/or other information may be obtained during the capture duration. An observed trajectory of the image capture device during the capture duration may be determined based on the position information and/or other information. The observed trajectory may reflect the positions of the image capture device at different moments within the capture duration. The observed trajectory may include a first portion corresponding to a first moment within the capture duration and a second portion corresponding to a second moment subsequent to the first moment within the capture duration. A capture trajectory of the

SMRH:486465515.2

-1-

46DG-274956

image capture device may be determined based on a subsequent portion of the observed trajectory and/or other information such that a portion of the capture trajectory corresponding to the first portion of the observed trajectory may be determined based on the second portion of the observed trajectory. The capture trajectory may have smoother changes in the positions of the image capture device than the observed trajectory. Orientations of the capture field of view for the images with respect to the optical field of view of the images may be determined based on the capture trajectory of the image capture device and/or other information. Video content may be generated based on visual content of the images within the capture field of view and/or other information.

(04) A system that stabilizes videos may include one or more electronic storages, one or more processors, and/or other components. An electronic storage may store image information, position information, information relating to an observed trajectory of an image capture device, information relating to a capture trajectory, information relating to an optical field of view, information relating to a capture field of view, information relating to video content, and/or other information. In some implementations, the system may include one or more image sensors, one or more position sensors, and/or other components.

(05) One or more components of the system may be carried by a housing, such as a housing of an image capture device. For example, the image sensor(s) and the position sensor(s) of the system may be carried by the housing. The housing may carry other components, such as the processor(s) and/or one or more optical elements. An optical

element may be configured to guide light within an optical field of view to an image sensor. The optical field of view may be greater than a capture field of view for generating video content. An image sensor may be configured to generate an image output signal based on light that becomes incident thereon during a capture duration. The image output signal may convey image information that defines images with the optical field of view.

(06) A position sensor may be configured to generate a position output signal based on positions of the housing during the capture duration. The position output signal may convey position information that characterizes positions of the housing at different moments within the capture duration. In some implementations, the position sensor may include one or more of a gyroscope, an accelerometer, and/or an inertial measurement unit. The position information may be determined independent of the image information.

(07) The processor(s) may be configured by machine-readable instructions. Executing the machine-readable instructions may cause the processor(s) to facilitate stabilizing videos. The machine-readable instructions may include one or more computer program components. The computer program components may include one or more of an observed trajectory component, a capture trajectory component, an orientation component, a generation component, and/or other computer program components.

(08) The observed trajectory component may be configured to determine an observed trajectory of the housing during the capture duration based on the position information

and/or other information. The observed trajectory may reflect positions of the housing at different moments within the capture duration. The positions of the housing may include rotational positions and/or translational positions of the housing. The observed trajectory may include a first portion corresponding to a first moment within the capture duration and a second portion corresponding to a second moment subsequent to the first moment within the capture duration.

(09) The capture trajectory component may be configured to determine a capture trajectory of the housing based on a subsequent portion of the observed trajectory and/or other information. The capture trajectory may be determined such that a portion of the capture trajectory corresponding to the first portion of the observed trajectory is determined based on the second portion of the observed trajectory. The capture trajectory may have smoother changes in the positions of the housing than the observed trajectory.

(10) In some implementations, the capture trajectory having smoother changes in the positions of the housing than the observed trajectory may be characterized by the capture trajectory having less jitters in the positions of the housing than the observed trajectory.

(11) In some implementations, the capture trajectory of the housing may be determined based on minimization of a rotational velocity of the housing and a rotational acceleration of the housing while respecting a set of constraints. The set of constraints may include a margin constraint, a trajectory constraint, a target constraint, and/or other constraints. The margin constraint may be determined based on a difference between

the optical field of view and the capture field of view, and/or other information. The trajectory constraint may be determined based on a subsequent portion of the observed trajectory and/or other information. The target constraint may be determined based on positions of a target in the images and/or other information.

(12) The orientation component may be configured to determine orientations of the capture field of view for the images with respect to the optical field of view of the images based on the capture trajectory of the housing and/or other information.

(13) The generation component may be configured to generate video content based on visual content of the images within the capture field of view and/or other information.

(14) These and other objects, features, and characteristics of the system and/or method disclosed herein, as well as the methods of operation and functions of the related elements of structure and the combination of parts and economies of manufacture, will become more apparent upon consideration of the following description and the appended claims with reference to the accompanying drawings, all of which form a part of this specification, wherein like reference numerals designate corresponding parts in the various figures. It is to be expressly understood, however, that the drawings are for the purpose of illustration and description only and are not intended as a definition of the limits of the invention. As used in the specification and in the claims, the singular form of "a," "an," and "the" include plural referents unless the context clearly dictates otherwise.

BRIEF DESCRIPTION OF THE DRAWINGS

- (15) FIG. 1 illustrates a system that stabilizes videos.
- (16) FIG. 2 illustrates a method for stabilizing videos.
- (17) FIG. 3 illustrates an example image capture device.
- (18) FIG. 4 illustrates an example observed trajectory.
- (19) FIG. 5A illustrates example predicted trajectories.
- (20) FIG. 5B illustrates an example smoothed trajectory.
- (21) FIG. 5C illustrates an example capture trajectory.
- (22) FIG. 6A illustrates example optical field of view and capture field of view.
- (23) FIG. 6B illustrates example orientations of a capture field of view with respect to an optical field of view.

DETAILED DESCRIPTION

(24) FIG. 1 illustrates a system 10 for stabilizing videos. The system 10 may include one or more of a processor 11, an interface 12 (e.g., bus, wireless interface), an electronic storage 13, and/or other components. In some implementations, the system 10 may include one or more image sensors, one or more position sensors, and/or other components. Images with an optical field of view may be captured by an image capture device during a capture duration. Image information defining the images, position information characterizing positions of the image capture device at different moments within the capture duration, and/or other information may be obtained during the capture

duration. Image information, position information, and/or other information may be obtained by the processor 11. An observed trajectory of the image capture device during the capture duration may be determined based on the position information and/or other information. The observed trajectory may reflect the positions of the image capture device at different moments within the capture duration. The observed trajectory may include a first portion corresponding to a first moment within the capture duration and a second portion corresponding to a second moment subsequent to the first moment within the capture duration. A capture trajectory of the image capture device may be determined based on a subsequent portion of the observed trajectory and/or other information such that a portion of the capture trajectory corresponding to the first portion of the observed trajectory may be determined based on the second portion of the observed trajectory. The capture trajectory may have smoother changes in the positions of the image capture device than the observed trajectory. Orientations of the capture field of view for the images with respect to the optical field of view of the images may be determined based on the capture trajectory of the image capture device and/or other information. Video content may be generated based on visual content of the images within the capture field of view and/or other information.

(25) The electronic storage 13 may be configured to include electronic storage medium that electronically stores information. The electronic storage 13 may store software algorithms, information determined by the processor 11, information received remotely, and/or other information that enables the system 10 to function properly. For example, the electronic storage 13 may store information relating to images, image

information, information relating to image capture device, information relating to optical element, information relating to image sensor, information relating to position sensor, position information, information relating to observed trajectory of the image capture device, information relating to capture trajectory of the image capture device, information relating to optical field of view, information relating to capture field of view, information relating to video content, and/or other information.

(26) The processor 11 may be configured to provide information processing capabilities in the system 10. As such, the processor 11 may comprise one or more of a digital processor, an analog processor, a digital circuit designed to process information, a central processing unit, a graphics processing unit, a microcontroller, an analog circuit designed to process information, a state machine, and/or other mechanisms for electronically processing information. The processor 11 may be configured to execute one or more machine-readable instructions 100 to facilitate stabilizing videos. The machine-readable instructions 100 may include one or more computer program components. The machine-readable instructions 100 may include one or more of an observed trajectory component 102, a capture trajectory component 104, an orientation component 106, a generation component 108, and/or other computer program components.

(27) Images with an optical field of view may be captured by an image capture device during a capture duration. An optical field of view of image may define a field of view of a scene captured within the image. A capture duration may be measured/defined in terms of time durations and/or frame numbers. For example, images may be captured

during a capture duration of 60 seconds, and/or from one point in time to another point in time. Images may be captured during a capture duration including capture of 1800 images. If the images are captured at 30 images/second, then the capture duration may correspond to 60 seconds. Other capture durations are contemplated.

(28) The system 10 may be remote from the image capture device or local to the image capture device. One or more portions of the image capture device may be remote from or a part of the system 10. One or more portions of the system 10 may be remote from or a part of the image capture device.

(29) For example, one or more components of the system 10 may be carried by a housing, such as a housing of an image capture device. For instance, image sensor(s) and position sensor(s) of the system 10 may be carried by the housing of the image capture device. The housing may carry other components, such as the processor 11 and/or one or more optical elements.

(30) An image capture device may refer to a device for recording visual information in the form of images, videos, and/or other media. An image capture device may be a standalone device (e.g., camera) or may be part of another device (e.g., part of a smartphone). FIG. 3 illustrates an example image capture device 302. The image capture device 302 may include a housing 312, and the housing 312 may carry (be attached to, support, hold, and/or otherwise carry) an optical element 304, an image sensor 306, a position sensor 308, a processor 310, and/or other components. Other configurations of image capture devices are contemplated.

(31) The optical element 304 may include instrument(s), tool(s), and/or medium that acts upon light passing through the instrument(s)/tool(s)/medium. For example, the optical element 304 may include one or more of lens, mirror, prism, and/or other optical elements. The optical element 304 may affect direction, deviation, and/or path of the light passing through the optical element 304. The optical element 304 may be configured to guide light within an optical field of view 305 to the image sensor 306. The optical field of view 305 may include the field of view of a scene that is within the field of view of the optical element 304 and/or the field of view of the scene that is delivered to the image sensor 306. For example, the optical element 304 may guide light within its field of view to the image sensor 306 or may guide light within a portion of its field of view to the image sensor 306. The optical field of view 305 may be greater than a capture field of view for generating video content.

(32) The image sensor 306 may include sensor(s) that converts received light into output signals. The output signals may include electrical signals. For example, the image sensor 306 may include one or more of a charge-coupled device sensor, an active pixel sensor, a complementary metal-oxide semiconductor sensor, an N-type metal-oxide-semiconductor sensor, and/or other image sensors. The image sensor 306 may generate output signals conveying information that defines one or more images (e.g., video frames of a video). For example, the image sensor 306 may be configured to generate an image output signal based on light that becomes incident thereon during a capture duration. The image output signal may convey image information that defines images with the optical field of view.

(33) The position sensor 308 may include sensor(s) that converts experienced positions/motions into output signals. The output signals may include electrical signals. For example, the position sensor 308 may refer to a set of position sensors, which may include one or more inertial measurement units, one or more accelerometers, one or more gyroscopes, and/or other position sensors. The position sensor 308 may generate output signals conveying information that characterizes positions of the position sensor 308 and/or device(s) carrying the position sensor 308. For example, the position sensor 308 may be configured to generate a position output signal based on positions of the housing/image capture device during the capture duration. The position output signal may convey position information that characterizes positions of the housing 312 at different moments (points in time, time durations) within the capture duration. The position information may characterize positions of the housing 312 based on specific translational and/or rotational positions of the housing 312 and/or based on changes in translational and/or rotational positions of the housing 312 as a function of progress through the capture duration. That is, the position information may characterize translational and/or rotational positions of the housing 312 and/or changes in translational and/or rotational positions (motion) of the housing 312 (e.g., direction, amount, velocity, acceleration) during the capture duration.

(34) The position information may be determined based on signals generated by the position sensor 308 and independent of the information/signals generated by the image sensor 306. That is, position information may be determined without using images/videos generated by the image sensor 306. Use of images/videos to determine

positions/motions of the housing 312/image capture device 302 may be computationally expensive in terms of processing power, processing time, and/or battery consumption.

Using the information/signals from the position sensor 308 to determine positions/motions of the housing 312/image capture device 302 may be computationally cheaper. That is, less processing power, processing time, and/or battery consumption may be required when positions/motions of the housing 312/image capture device 302 are determined from the information/signals from the position sensor 308 than the information/signals from the image sensor 306. The position information determined independent of the image information may be used to determine the trajectory of the housing 312/image capture device 302 during the capture duration.

(35) The processor 310 may include one or more processors (logic circuitry) that provide information processing capabilities in the image capture device 302. The processor 310 may provide one or more computing functions for the image capture device 302. The processor 310 may operate/send command signals to one or more components of the image capture device 302 to operate the image capture device 302. For example, the processor 310 may facilitate operation of the image capture device 302 in capturing image(s) and/or video(s), facilitate operation of the optical element 304 (e.g., change how light is guided by the optical element 304), and/or facilitate operation of the image sensor 306 (e.g., change how the received light is converted into information that defines images/videos and/or how the images/videos are post-processed after capture). The processor 310 may obtain information from the image sensor 306 and/or the position sensor 308, and/or facilitate transfer of information from

the image sensor 306 and/or the position sensor 308 to another device/component.

The processor 310 may be remote from the processor 11 or local to the processor 11.

One or more portions of the processor 310 may be part of the processor 11 and/or one or more portions of the processor 10 may be part of the processor 310.

(36) Referring back to FIG. 1, the processor 11 (or one or more components of the processor 11) may be configured to obtain information to facilitate stabilizing videos. Obtaining information may include one or more of accessing, acquiring, analyzing, determining, examining, identifying, loading, locating, opening, receiving, retrieving, reviewing, storing, and/or otherwise obtaining the information. The processor 11 may obtain information from one or more locations. For example, the processor 11 may obtain information from a storage location, such as the electronic storage 13, electronic storage of information and/or signals generated by one or more sensors, electronic storage of a device accessible via a network, and/or other locations. The processor 11 may obtain information from one or more hardware components (e.g., an image sensor, a position sensor) and/or one or more software components (e.g., software running on a computing device).

(37) For example, the processor 11 (or one or more components of the processor 11) may obtain image information defining images with an optical field of view, position information characterizing positions of an image capture device at different moments within a capture duration, and/or other information. One or more information may be obtained during acquisition of the images and/or after acquisition of the images by an image capture device. For example, image information, position information, and/or

other information may be obtained while the images are being captured by an image capture device and/or after the images have been captured and stored in memory (e.g., the electronic storage 13).

(38) The observed trajectory component 102 may be configured to determine an observed trajectory of the image capture device/housing of the image capture device during the capture duration based on the position information and/or other information. The determination of the observed trajectory may be referred to as trajectory generation/observed trajectory generation. The observed trajectory may refer to one or more paths and/or progression followed by the image capture device/housing during the capture duration. The observed trajectory may reflect positions of the image capture device/housing of the image capture device at different moments within the capture duration. The positions of the image capture device/housing of the image capture device may include rotational positions (e.g., rotations about one or more axis of the image capture device) and/or translational positions of the image capture device/housing of the image capture device. For example, the observed trajectory component 102 may determine an observed trajectory of the image capture device/housing of the image capture device during the capture duration based on the position information characterizing specific translational and/or rotational positions of the image capture device/housing and/or changes in translational and/or rotational positions of the image capture device/housing as a function of progress through the capture duration.

(39) The observed trajectory may include different portions corresponding to different moments within the capture duration. For example, the observed trajectory may include a first portion corresponding to a first moment within the capture duration and a second portion corresponding to a second moment within the capture duration. The second moment may be subsequent to the first moment within the capture duration.

(40) FIG. 4 illustrates an example observed trajectory of an image capture device. The observed trajectory may include an observed yaw trajectory 400 of the image capture device. The observed yaw trajectory 400 may reflect yaw angle positions (e.g., rotational positions defined with respect to a yaw axis, rotation to the left or right) of the image capture device/housing of the image capture device at different moments within the capture duration. The observed yaw trajectory 400 may show that the image capture device was rotated in a negative yaw direction, rotated in a positive yaw direction, rotated back in the forward configuration, then rotated in the negative yaw direction. For example, the image capture device, during capture of images, may have been rotated to the right, then to the left, to the front, then to the right. Other types of observed trajectory (e.g., pitch trajectory, roll trajectory, translational trajectory) are contemplated.

(41) Generating a video based on the images captured along the observed yaw trajectory 400 may be undesirable. For example, generating a video based on the images captured along the observed yaw trajectory 400 may result in a video with footage that is shaky and/or that appears to include unintended camera motion. For instance, sharp/quick changes in the yaw angle positions of the image capture device

may result in abrupt changes in the direction of visuals within the video (e.g., quick left or right camera motion). Multiple changes in the yaw angle positions of the image capture device may result in a footage that is changing the direction of view (e.g., to the right, to the left, to the front, to the right).

(42) Videos generated from such images may be stabilized to provide a smoother view of the captured content. Stabilization of such videos may include using smaller visual content portions of the images to provide a punch-out view of the images that creates a more stable view than when generating videos by using the entire visual content of the images. For example, such stabilization may be provided by creating a stabilized trajectory over the capture duration and determining the punch-outs from the images based on the stabilized trajectory. A punch-out of an image may refer to one or more portions of the image that is used for presentation, such as a cut-out of the image or a cropped portion of the image. A punch-out of an image may include one or more visual portions of the image presented on a display and/or one or more visual portion portions of the image used to generate video frames of video content. However, some stabilization techniques may not preserve the intent of the user that captured the images.

(43) For example, videos may be stabilized by predicting positions/motions of the camera based on past positions/motions. For example, when attempting to determine the location and/or shape of the punch-out for an image captured at a given moment within the capture duration, the positions/motions of the image capture device preceding that moment may be used to determine how to position/shape the punch-out to create a

stabilized view. Such use of “past” position/motion information may conflict with the motion intended by the user of the image capture device.

(44) For example, in the observed yaw trajectory 400, the rotation of the image capture device to the right, then to the left, and then to the front may be the result of the image capture device being unintentionally rotated to the right, the user overcorrecting the rotation to the left, and then rotating the image capture device to the right to a front-facing direction. Determining punch-out of the images using “past” position/motion information may result in a predicted trajectory as shown in FIG. 5A.

(45) For example, based on the rotation of the image capture device to the right during duration A 502, a predicted yaw trajectory A 512, which continues the rotation to the right, may be predicted for duration B 504. Based on smaller rotation of the image capture device to the right during duration D 506, a predicted yaw trajectory B 514, which continues the smaller rotation to the right, may be predicted for duration D 508. The predicted yaw trajectory A 512 may be in opposite direction of the actual motion of the image capture device during duration B 504, and the predicted yaw trajectory B 514 may be off from the observed yaw trajectory by a large margin. Such discrepancy between the observed yaw trajectory and the predicted yaw trajectory may result in the images not including enough visual information (e.g., pixels) to account for the attempted stabilized and/or the punch-out location/shape.

(46) As another example, videos may be stabilized by smoothing the observed changes in positions/motions of the image capture device. For instance, a low-pass

filter may be applied to the observed trajectory to smooth out the abrupt rotational and/or translational changes in the positions/motions of the image capture device.

(47) For example, as shown in FIG. 5B, by applying a low-pass filter to the observed yaw trajectory 400, a smoothed yaw trajectory 516 may be determined. The smoothed yaw trajectory 516 may have smoother changes in the positions/motions of the image capture device than the observed yaw trajectory. However, such smooth may not take into account how the positions/motions changes during the capture duration and may not preserve the intent of the user that captured the images. For example, even though a video generated from the smoothed yaw trajectory 516 may not include abrupt changes in the direction of visuals within the video, the video may still include footage that is changing the direction of view to the right, then to left, to the front, and then to the right in a non-continuous motion (e.g., rotating to the right to a certain angle, holding that position for a time, then rotating to the right, then back a little to the left, and then to the right again).

(48) The capture trajectory component 104 may be configured to determine a capture trajectory of the image capture device/housing of the image capture device based on a subsequent portion of the observed trajectory and/or other information. The determination of the capture trajectory may be referred to as trajectory generation/capture trajectory generation. The capture trajectory may refer to one or more paths and/or progression to be used in determining which portions of the visual content of the images (punch-outs) may be used to generate a video. The capture trajectory may reflect actual and/or virtual positions of the image capture device/housing

of the image capture device at different moments within the capture duration. An actual position may refer to a position that was taken by the image capture device/housing of the image capture. A virtual position may refer to a position that was not taken by the image capture device/housing of the image capture. A virtual position may be offset (rotationally and/or translationally) from the actual position of the image capture device/housing of the image capture device. The capture trajectory may have smoother changes in the positions of the image capture device/housing of the image capture device than the observed trajectory. That is, the capture trajectory may have less jitters (slight irregular movement/variation), less abrupt changes, and/or less discontinuous changes in the positions of the image capture device/housing of the image capture device than the observed trajectory.

(49) Determining a capture trajectory (capture trajectory generation) based on a subsequent portion of an observed trajectory may include determining a portion of the capture trajectory corresponding to a given moment within the capture duration based on one or more portions of the observed trajectory corresponding to one or more subsequent moments (moment(s) past the given moment) within the capture duration. That is, the capture trajectory component 104 may “look ahead” in time to determine a portion of the capture trajectory. A look ahead may include use of one or more subsequent portions of an observed trajectory to determine a preceding portion of the capture trajectory. Such generation of trajectory may be referred to as a look head trajectory generation. A subsequent moment within the duration may be adjacent to the given moment or not adjacent to the given moment. Using the subsequent portion(s) of

the observed trajectory may enable the capture trajectory component 104 to determine a capture trajectory that preserves a user's intended motion for the image capture device. The user's intended motion may refer to the motion of the image capture device that the user planned/meant to carry out.

(50) The positions/motions of the image capture device in the "future" may be analyzed (look ahead) to determine whether particular position(s)/motion(s) of the image capture device at a moment is an intended motion or an unintended motion (e.g., shaking due to vibration, rotation due to bumping/mishandling the image capture device). For example, when determining the capture trajectory for a moment (e.g., corresponding to the 1000th video frame) within the capture duration, the position(s)/motion(s) of the image capture device for a duration subsequent to the moment (e.g., corresponding to the next 30 video frames) may be analyzed to determine whether the position/motion of the image capture device at the moment was intended or not. In some implementations, the capture trajectory component 104 may determine a capture trajectory of the image capture device/housing of the image capture device further based on one or more prior portions of the observed trajectory. The past position/motion information of the image capture device may supply context for the intended motion.

(51) FIG. 5C illustrates an example capture trajectory determined by the capture trajectory component 104. The capture trajectory may include a capture yaw trajectory 532. The capture yaw trajectory 532 may reflect yaw angle positions (e.g., rotational positions defined with respect to a yaw axis, rotation to the left or right) of the image

capture device/housing of the image capture device to be used in determining which portions of the visual content of the images (punch-outs) may be used to generate a video. For example, the capture yaw trajectory 532 may include a zero-rotation about the yaw axis (front-direction) for durations 522, 524, 526 and then a smooth rotation to the right for durations 528, 530. Other types of capture trajectory (e.g., capture pitch trajectory, capture roll trajectory, capture translational trajectory) are contemplated.

(52) The capture yaw trajectory 532 may be determined such that a portion of the capture yaw trajectory 532 corresponding to the portion of the observed yaw trajectory 400 is determined based on a subsequent portion of the observed yaw trajectory 400. For example, the portion of the capture yaw trajectory 532 for one or more portions of the duration A 522 may be determined based on the portion(s) of the observed yaw trajectory 400 within the duration B 524 and/or the duration C 526 (look ahead to the duration B 524 and/or the duration C 526). The portion(s) of the observed yaw trajectory 400 for the duration B 524 and/or the duration C 526 may be used to determine in what direction and/or by what amount the capture yaw trajectory 532 for portion(s) of the duration A 522 may differ from the observed yaw trajectory 400. The capture yaw trajectory 532 may be determined based on subsequent portion(s) of the observed trajectory 400 such that the capture yaw trajectory 532 preserves a user's intended motion for the image capture device. For example, based on the subsequent portion(s) (look ahead) of the observed trajectory 400, the capture trajectory component 104 may determine that the rotation of the image capture device to the right and left during durations 522, 524 were unintended motions (e.g., the image capture device

being unintentionally rotated to the right and the user overcorrecting the rotation to the left), and may determine the capture yaw trajectory 532 to be directed to the front during durations 522, 524. Based on the subsequent portion(s) (look ahead) of the observed trajectory 400, the capture trajectory component 104 may determine that the staggered rotation of the image capture device to the right during durations 528, 530 included unintended motions (non-continuous rotation to the right), and may determine the capture yaw trajectory 532 to include continuous rotation to the right during durations 528, 530. Other determinations of capture trajectory are contemplated.

(53) In some implementations, the capture trajectory may be determined based on minimization of a rotational velocity of the image capture device/housing of the image capture device and a rotational acceleration of the image capture device/housing of the image capture device while respecting a set of constraints. For example, the capture trajectory may be determined by generating a smooth path that respects the set of constraints, rather than by modifying the observed trajectory. For instance, a smooth path defining yaw angle position, pitch angle position, and/or roll angle position may be generated by finding a path of the image capture device/housing of the image capture device that minimizes a combination of a time derivative, a second time derivative, and/or other time derivative(s) of the yaw angle position, pitch angle position, and/or roll angle position while respecting the set of constraints:

$$(54) \quad \min \left(\left\| \frac{d}{dt} \theta \right\|^2 + \left\| \frac{d^2}{dt^2} \theta \right\|^2 + \dots \right)$$

(55) In some implementations, one or more of the portions of the minimization calculation may be changed. For example, one or more portions of the minimization calculation (e.g., the first time derivative) may be changed to have a greater affect or a lesser affect than other portion(s) of the minimization calculation (e.g., the second time derivative), and/or other factors may be introduced into the calculation.

(56) In some implementations, information on high frequencies (jitters) of image capture may be used to improve visual characteristics of generated video content. Certain portion of high frequencies in the input may be kept based on the image capture configuration, such as exposure start time and exposure duration time, the position information (e.g., position sensor readings), and/or other information. For example, motion of the image capture device/image sensor during a frame exposure may be analyzed and used to generate/modify a capture trajectory that minimizes inter-frame motion (e.g., smooths inter-frame motion) while preserving the intra-frame motion, which may contain the high frequencies. This may provide for improved visual characteristics of the generated video content, such as by compensating for motion blur and/or lowlight image capture conditions.

(57) For example, image capture may not happen instantaneously. Rather, it may take a certain amount of time for image sensor pixel sites to gather light. This may lead to splitting of the image sensor motion in time in two phases: inter-frame motion which may not be captured and may be suppressed, and intra-frame motion which may be “encrusted” in the image and may not be removed. Better visual characteristics (e.g., impression) may be provided when the intra-frame motion is taken into account so that

the capture trajectory is smooth for inter-frame phases while corresponding to the original motion for intra-frame phases. That is, the capture trajectory may move/follow in the same direction and with the same speed as during the frame exposure phase, so that its motion is aligned with the motion blur in the image.

(58) The set of constraints may include one or more constraints that provide limitations/restrictions/rules on how the smooth path is generated. For example, the set of constraints may include a margin constraint that provide limitation(s)/restriction(s)/rule(s) on how far from the observed trajectory the smooth path may be generated. The margin constraint may be determined based on a difference between the optical field of view and the capture field of view, and/or other information. The optical field of view may refer to the field of view of a scene captured within the image. That is, the optical field of view may refer to the spatial extent/angle of the scene captured within the image. The capture field of view may refer to the field of view used to generate a video based on visual content of the images. That is, a video may be generated based on visual content of the images within the capture field of view. The capture field of view may be defined in terms of shape and/or size.

(59) For example, FIG. 6A illustrates an example optical field of view 602 and an example capture field of view 604 for an image A 600. The image A 600 may include capture of a scene within an angle defined by the optical field of view 602. The capture field of view 604 may provide a punch-out of the image A 600 to be used for video generation. The amount and/or direction from which the smooth path may deviate from the observed trajectory may depend on the difference between the optical field of view

602 and the capture field of view 604. The difference between the optical field of view 602 and the capture field of view 604 (e.g., 10%) may define a margin 606 within which the capture field of view 604 may move with respect to the optical field of view 602. For example, referring to FIG. 6B, a capture field of view 614 may be rotated with respect to the optical field of view 612 while not going beyond the pixels captured within an image B 610, and a capture field of view 624 may be laterally moved with respect to a optical field of view 622 while not going beyond the pixels captured within an image C 620.

(60) Larger difference between the optical field of view 602 and the capture field of view 604 may enable larger movement of the capture field of view 604 with respect to the optical field of view 602 while smaller difference between the optical field of view 602 and the capture field of view 604 may enable smaller movement of the capture field of view 604 with respect to the optical field of view 602. However, the larger margin 606 may result in waste of pixel space and computing resources (e.g., processor power and/or battery consumption to capture images with larger optical field of view than needed to generate a video).

(61) In some implementations, the set of constraints may include a trajectory constraint that provide limitation(s)/restriction(s)/rule(s) on how the smooth path may be generated based on subsequent portions of the observed trajectory. The trajectory constraint may be determined based on a subsequent portion of the observed trajectory and/or other information. That is, the trajectory constraint may include one or more constraints relating to shape of the observed trajectory in the “future.” The trajectory

constraint may preserve the intended motion for the image capture device in the generated path.

(62) In some implementations, the set of constraints may include a target constraint that provide limitation(s)/restriction(s)/rule(s) on how the smooth path may be generated based on a target within the images. A target may refer to a person, an object, and/or a thing that may be selected for inclusion in a video. For example, the images captured by the image capture device may include one or more view of a person (e.g., a person of interest) and a user may wish to create a video that includes the person. The target constraint may include one or more constraints relating to the location of the target within the images such that the images are stabilized around the location of the target within the images. That is, the target constraint may affect the generation of the smooth path so that the target is within one or more of the punch-outs for the images. Other constraints are contemplated.

(63) The orientation component 106 may be configured to determine orientations of the capture field of view for the images with respect to the optical field of view of the images based on the capture trajectory of the image capture device/housing of the image capture device and/or other information. The orientations of the capture field of view for the images with respect to the optical field of view of the images may determine which portions of the visual content of the images (punch-outs) may be used to generate a video. That is, the orientation component 106 may determine how the punch-outs for the images may be oriented with respect to the capture field of view for the images. The capture trajectory may be used to determine how much and in what

direction the capture field of view is rotated with respect to the optical field of view. In some implementations, the orientation component 106 may determine how the punch-outs for the images may be laterally and/or vertically positioned with respect to the capture field of view for the images. The capture trajectory may determine how much and in what direction the capture field of view is rotated with respect to the optical field of view. The capture trajectory may be used to determine how much and in what direction the capture field of view is laterally/vertically positioned with respect to the optical field of view.

(64) For example, referring to FIG. 6B, the orientation component 106 may determine the orientation of the capture field of view 614 with respect to the optical field of view 612 for the image B 610 based on the capture trajectory at a particular moment. The capture field of view 614 may be oriented with respect to the optical field of view 612 to provide a punch-out of the image B 610 that is stable with respect to a prior and/or next punch-out of the images (e.g., stable with the punch-out of the image A 600 using the capture field of view 604 for the image A 600, shown in FIG. 6A). The orientation component 106 may determine the orientation of the capture field of view 624 with respect to the optical field of view 622 for the image C 620 based on the capture trajectory at a particular moment. The capture field of view 624 may be oriented with respect to the optical field of view 622 to provide a punch-out of the image C 620 that is stable with respect to a prior and/or next punch-out of the images (e.g., stable with the punch-out of the image B 610 using the capture field of view 614 for the image B 610).

(65) The generation component 108 may be configured to generate video content based on visual content of the images within the capture field of view and/or other information. Video content may refer media content that may be consumed as one or more videos/video clips. Video content may include one or more videos/video clips stored in one or more formats/containers, and/or other video content. A format may refer to one or more ways in which the information defining video content is arranged/laid out (*e.g.*, file format). A container may refer to one or more ways in which information defining video content is arranged/laid out in association with other information (*e.g.*, wrapper format). Video content may define visual content viewable as a function of progress through a progress length of the video content. Video content may include video frames that define visual content. That is, visual content of the video content may be included within video frames of the video content.

(66) The video frames of the video content may be determined based on visual content of the images within the capture field of view and/or other information. The video frames of the video content may be determined based on a punch-out of the images in accordance with the capture trajectory, the capture field of view, and/or other information. For example, referring to FIGS. 6A-6B, the video frames of the video content may be determined based on the visual content of the image A 600 within the capture field of view 604, the visual content of the image B 610 within the capture field of view 614, the visual content of the image C 620 within the capture field of view 624, and/or other information. Such determination of visual content of images for inclusion in video content may effectuate stabilization of the video content.

(67) In some implementations, the video frames of the video content may be determined based on warping of images (e.g., one or more portions of the visual content of the images). The warping of the images may provide for different perspectives of content captured within the images, with the different perspectives corresponding to how the content would have look had the images been captured from the image capture device on the capture trajectory.

(68) In some implementations, one or more of the images and/or one or more of the portions of images used to generate video content may be stored in a buffer (e.g., 1s buffer). The buffer may be used to store images/portions of images including visual content that will be included in the visual content and/or images/portions of images including visual content that will be transformed (e.g., warped) for inclusion in the visual content. The buffer may be used to store image information, position information, and/or other information for look ahead and/or trajectory generation. For example, the buffer may be used to store images for which trajectory generation is being performed using look ahead. After corresponding portion of the capture trajectory is generated, the relevant portions of the images (visual content of the images within the capture field of view) in the buffer may be used to generate the video content.

(69) The video content generated by the generation component 108 may be defined by video information. Video information defining video content may define an encoded version/copy of the video content and/or instructions for rendering the video content. For example, the video information may define an encoded version/copy of the video content, and the video information (e.g., video file) may be opened in a video player for

presentation of the video content. The video information may define instructions to render the video content for presentation. For example, the video information may define a director track that includes information as to which visual portions of the images should be included within the presentation of the video content. The director track may include information on the location and/or shape of the punch-out of images to be used to as a function progress through the video content to provide a stabilized view of the images. A video player may use the director track to retrieve the relevant visual portions of the images when the video content is opened/to be presented.

(70) The generation component 108 may be configured effectuate storage of the video information and/or other information in one or more storage media. For example, the video information may be stored in the electronic storage 13, remote storage locations (storage media located at/accessible through a server), and/or other locations. In some implementations, the generation component 108 may effectuate storage of the video information through one or more intermediary devices. For example, the processor 11 may be located within a computing device without a connection to the storage device (e.g., the computing device lacks WiFi/cellular connection to the storage device). The generation component 108 may effectuate storage of the video information through another device that has the necessary connection (e.g., the computing device using a WiFi/cellular connection of a paired mobile device, such as a smartphone, tablet, laptop, to store information in one or more storage media). Other storage locations for and storage of the video information are contemplated.

(71) While the description herein may be directed to images and videos, one or more other implementations of the system/method described herein may be configured for other types media content. Other types of media content may include one or more of audio content (e.g., music, podcasts, audiobooks, and/or other audio content), multimedia presentations, images, slideshows, visual content (e.g., one or more images and/or videos), and/or other media content.

(72) Implementations of the disclosure may be made in hardware, firmware, software, or any suitable combination thereof. Aspects of the disclosure may be implemented as instructions stored on a machine-readable medium, which may be read and executed by one or more processors. A machine-readable medium may include any mechanism for storing or transmitting information in a form readable by a machine (e.g., a computing device). For example, a tangible computer-readable storage medium may include read-only memory, random access memory, magnetic disk storage media, optical storage media, flash memory devices, and others, and a machine-readable transmission media may include forms of propagated signals, such as carrier waves, infrared signals, digital signals, and others. Firmware, software, routines, or instructions may be described herein in terms of specific exemplary aspects and implementations of the disclosure, and performing certain actions.

(73) In some implementations, some or all of the functionalities attributed herein to the system 10 may be provided by external resources not included in the system 10. External resources may include hosts/sources of information, computing, and/or

processing and/or other providers of information, computing, and/or processing outside of the system 10.

(74) Although the processor 11 and the electronic storage 13 are shown to be connected to the interface 12 in FIG. 1, any communication medium may be used to facilitate interaction between any components of the system 10. One or more components of the system 10 may communicate with each other through hard-wired communication, wireless communication, or both. For example, one or more components of the system 10 may communicate with each other through a network. For example, the processor 11 may wirelessly communicate with the electronic storage 13. By way of non-limiting example, wireless communication may include one or more of radio communication, Bluetooth communication, Wi-Fi communication, cellular communication, infrared communication, Li-Fi communication, or other wireless communication. Other types of communications are contemplated by the present disclosure.

(75) Although the processor 11 is shown in FIG. 1 as a single entity, this is for illustrative purposes only. In some implementations, the processor 11 may comprise a plurality of processing units. These processing units may be physically located within the same device, or the processor 11 may represent processing functionality of a plurality of devices operating in coordination. The processor 11 may be configured to execute one or more components by software; hardware; firmware; some combination of software, hardware, and/or firmware; and/or other mechanisms for configuring processing capabilities on the processor 11.

(76) It should be appreciated that although computer components are illustrated in FIG. 1 as being co-located within a single processing unit, in implementations in which processor 11 comprises multiple processing units, one or more of computer program components may be located remotely from the other computer program components.

(77) While computer program components are described herein as being implemented via processor 11 through machine-readable instructions 100, this is merely for ease of reference and is not meant to be limiting. In some implementations, one or more functions of computer program components described herein may be implemented via hardware (e.g., dedicated chip, field-programmable gate array) rather than software. One or more functions of computer program components described herein may be software-implemented, hardware-implemented, or software and hardware-implemented

(78) The description of the functionality provided by the different computer program components described herein is for illustrative purposes, and is not intended to be limiting, as any of computer program components may provide more or less functionality than is described. For example, one or more of computer program components may be eliminated, and some or all of its functionality may be provided by other computer program components. As another example, processor 11 may be configured to execute one or more additional computer program components that may perform some or all of the functionality attributed to one or more of computer program components described herein.

(79) The electronic storage media of the electronic storage 13 may be provided integrally (i.e., substantially non-removable) with one or more components of the system 10 and/or removable storage that is connectable to one or more components of the system 10 via, for example, a port (e.g., a USB port, a Firewire port, etc.) or a drive (e.g., a disk drive, etc.). The electronic storage 13 may include one or more of optically readable storage media (e.g., optical disks, etc.), magnetically readable storage media (e.g., magnetic tape, magnetic hard drive, floppy drive, etc.), electrical charge-based storage media (e.g., EPROM, EEPROM, RAM, etc.), solid-state storage media (e.g., flash drive, etc.), and/or other electronically readable storage media. The electronic storage 13 may be a separate component within the system 10, or the electronic storage 13 may be provided integrally with one or more other components of the system 10 (e.g., the processor 11). Although the electronic storage 13 is shown in FIG. 1 as a single entity, this is for illustrative purposes only. In some implementations, the electronic storage 13 may comprise a plurality of storage units. These storage units may be physically located within the same device, or the electronic storage 13 may represent storage functionality of a plurality of devices operating in coordination.

(80) FIG. 2 illustrates method 200 for stabilizing videos. The operations of method 200 presented below are intended to be illustrative. In some implementations, method 200 may be accomplished with one or more additional operations not described, and/or without one or more of the operations discussed. In some implementations, two or more of the operations may occur substantially simultaneously.

(81) In some implementations, method 200 may be implemented in one or more processing devices (e.g., a digital processor, an analog processor, a digital circuit designed to process information, a central processing unit, a graphics processing unit, a microcontroller, an analog circuit designed to process information, a state machine, and/or other mechanisms for electronically processing information). The one or more processing devices may include one or more devices executing some or all of the operation of method 200 in response to instructions stored electronically on one or more electronic storage mediums. The one or more processing devices may include one or more devices configured through hardware, firmware, and/or software to be specifically designed for execution of one or more of the operation of method 200.

(82) Referring to FIG. 2 and method 200, at operation 201, an image output signal may be generated. The image output signal may convey image information that defines images with an optical field of view. In some implementation, operation 201 may be performed by a component the same as or similar to the image sensor 306 (Shown in FIG. 3 and described herein).

(83) At operation 202, a position output signal may be generated. The position output signal may convey position information that characterizes positions of an image capture device at different moments within a capture duration. In some implementation, operation 202 may be performed by a processor component the same as or similar to the positions sensor 308 (Shown in FIG. 3 and described herein).

(84) At operation 203, an observed trajectory of the image capture device during the capture duration may be determined based on the position information. In some

implementation, operation 203 may be performed by a processor component the same as or similar to the observed trajectory component 102 (Shown in FIG. 1 and described herein).

(85) At operation 204, a capture trajectory of the image capture device may be determined based on a subsequent portion of the observed trajectory. In some implementation, operation 204 may be performed by a processor component the same as or similar to the capture trajectory component 104 (Shown in FIG. 1 and described herein).

(86) At operation 205, orientations of a capture field of view for the image may be determined with respect to the optical field of view based on the capture trajectory. In some implementation, operation 205 may be performed by a processor component the same as or similar to the orientation component 106 (Shown in FIG. 1 and described herein).

(87) At operation 206, video content may be generated based on visual content of the images within the capture field of view. In some implementation, operation 206 may be performed by a processor component the same as or similar to the generation component 108 (Shown in FIG. 1 and described herein).

(88) Although the system(s) and/or method(s) of this disclosure have been described in detail for the purpose of illustration based on what is currently considered to be the most practical and preferred implementations, it is to be understood that such detail is solely for that purpose and that the disclosure is not limited to the disclosed

implementations, but, on the contrary, is intended to cover modifications and equivalent arrangements that are within the spirit and scope of the appended claims. For example, it is to be understood that the present disclosure contemplates that, to the extent possible, one or more features of any implementation can be combined with one or more features of any other implementation.

What is claimed is:

1. A system that stabilizes videos, the system comprising:
 - a housing;
 - an optical element carried by the housing and configured to guide light within an optical field of view to an image sensor, the optical field of view being greater than a capture field of view for generating video content;
 - the image sensor carried by the housing and configured to generate an image output signal based on light that becomes incident thereon during a capture duration, the image output signal conveying image information that defines images with the optical field of view;
 - a position sensor carried by the housing and configured to generate a position output signal based on positions of the housing during the capture duration, the position output signal conveying position information that characterizes positions of the housing at different moments within the capture duration; and
 - one or more physical processors configured by machine-readable instructions to:
 - determine an observed trajectory of the housing during the capture duration based on the position information, the observed trajectory reflecting the positions of the housing at the different moments within the capture duration, the observed trajectory including a first portion corresponding to a first moment within the capture duration and a second portion corresponding to a second moment subsequent to the first moment within the capture duration;

determine a capture trajectory of the housing based on a subsequent portion of the observed trajectory such that a portion of the capture trajectory corresponding to the first portion of the observed trajectory is determined based on the second portion of the observed trajectory, the capture trajectory having smoother changes in the positions of the housing than the observed trajectory;

determine orientations of the capture field of view for the images with respect to the optical field of view of the images based on the capture trajectory of the housing; and

generate the video content based on visual content of the images within the capture field of view.

2. The system of claim 1, wherein the positions of the housing include rotational positions of the housing.
3. The system of claim 2, wherein the positions of the housing include translational positions of the housing.
4. The system of claim 1, wherein the position sensor includes a gyroscope, an accelerometer, and/or an inertial measurement unit, and the position information is determined independent of the image information.

5. The system of claim 1, wherein the capture trajectory having smoother changes in the positions of the housing than the observed trajectory is characterized by the capture trajectory having less jitters in the positions of the housing than the observed trajectory.

6. The system of claim 1, wherein the capture trajectory of the housing is determined based on minimization of a rotational velocity of the housing and a rotational acceleration of the housing while respecting a set of constraints.

7. The system of claim 6, wherein the set of constraints includes a margin constraint determined based on a difference between the optical field of view and the capture field of view.

8. The system of claim 7, wherein the set of constraints includes a trajectory constraint determined based on the subsequent portion of the observed trajectory.

9. The system of claim 8, wherein the set of constraints includes a target constraint determined based on positions of a target in the images.

10. A method for stabilizing videos, the method performed by a system including a housing and one or more processors, the housing carrying an optical element, an image sensor, and a position sensor, the optical element configured to guide light within an

optical field of view to the image sensor, the optical field of view being greater than a capture field of view for generating video content, the method comprising:

generating, by the image sensor, an image output signal based on light that becomes incident thereon during a capture duration, the image output signal conveying image information that defines images with the optical field of view;

generating, by the position sensor, a position output signal based on positions of the housing during the capture duration, the position output signal conveying position information that characterizes positions of the housing at different moments within the capture duration;

determining, by the one or more processors, an observed trajectory of the housing during the capture duration based on the position information, the observed trajectory reflecting the positions of the housing at the different moments within the capture duration, the observed trajectory including a first portion corresponding to a first moment within the capture duration and a second portion corresponding to a second moment subsequent to the first moment within the capture duration;

determining, by the one or more processors, a capture trajectory of the housing based on a subsequent portion of the observed trajectory such that a portion of the capture trajectory corresponding to the first portion of the observed trajectory is determined based on the second portion of the observed trajectory, the capture trajectory having smoother changes in the positions of the housing than the observed trajectory;

determining, by the one or more processors, orientations of the capture field of view for the images with respect to the optical field of view of the images based on the capture trajectory of the housing; and

generating, by the one or more processors, the video content based on visual content of the images within the capture field of view.

11. The method of claim 10, wherein the positions of the housing include rotational positions of the housing.

12. The method of claim 11, wherein the positions of the housing include translational positions of the housing.

13. The method of claim 10, wherein the position sensor includes a gyroscope, an accelerometer, and/or an inertial measurement unit, and the position information is determined independent of the image information.

14. The method of claim 10, wherein the capture trajectory having smoother changes in the positions of the housing than the observed trajectory is characterized by the capture trajectory having less jitters in the positions of the housing than the observed trajectory.

15. The method of claim 10, wherein the capture trajectory of the housing is determined based on minimization of a rotational velocity of the housing and a rotational acceleration of the housing while respecting a set of constraints.

16. The method of claim 15, wherein the set of constraints includes a margin constraint determined based on a difference between the optical field of view and the capture field of view.

17. The method of claim 16, wherein the set of constraints includes a trajectory constraint determined based on the subsequent portion of the observed trajectory.

18. The method of claim 17, wherein the set of constraints includes a target constraint determined based on positions of a target in the images.

19. A system that stabilizes videos, the system comprising:
a housing;
an optical element carried by the housing and configured to guide light within an optical field of view to an image sensor, the optical field of view being greater than a capture field of view for generating video content;
the image sensor carried by the housing and configured to generate an image output signal based on light that becomes incident thereon during a capture duration,

the image output signal conveying image information that defines images with the optical field of view;

a position sensor carried by the housing and configured to generate a position output signal based on positions of the housing during the capture duration, the positions of the housing including rotational positions and/or translational positions of the housing, the position output signal conveying position information that characterizes positions of the housing at different moments within the capture duration, the position information determined independent of the image information; and

one or more physical processors configured by machine-readable instructions to:

determine an observed trajectory of the housing during the capture duration based on the position information, the observed trajectory reflecting the positions of the housing at the different moments within the capture duration, the observed trajectory including a first portion corresponding to a first moment within the capture duration and a second portion corresponding to a second moment subsequent to the first moment within the capture duration;

determine a capture trajectory of the housing based on a subsequent portion of the observed trajectory such that a portion of the capture trajectory corresponding to the first portion of the observed trajectory is determined based on the second portion of the observed trajectory, the capture trajectory having smoother changes in the positions of the housing than the observed trajectory;

determine orientations of the capture field of view for the images with respect to the optical field of view of the images based on the capture trajectory of the housing; and

generate the video content based on visual content of the images within the capture field of view.

20. The system of claim 19, wherein the capture trajectory of the housing is determined based on minimization of a rotational velocity of the housing and a rotational acceleration of the housing while respecting a set of constraints, the set of constraints including a margin constraint determined based on a difference between the optical field of view and the capture field of view and a trajectory constraint determined based on the subsequent portion of the observed trajectory.

ABSTRACT

Images with an optical field of view may be captured by an image capture device. An observed trajectory of the image capture device reflecting the positions of the image capture device at different moments may be determined. A capture trajectory of the image capture device may be determined based on a subsequent portion of the observed trajectory such that a portion of the capture trajectory corresponding to the one portion of the observed trajectory may be determined based on a subsequent portion of the observed trajectory. Orientations of a capture field of view for the images with respect to the optical field of view of the images may be determined based on the capture trajectory. Video content may be generated based on visual content of the images within the capture field of view.

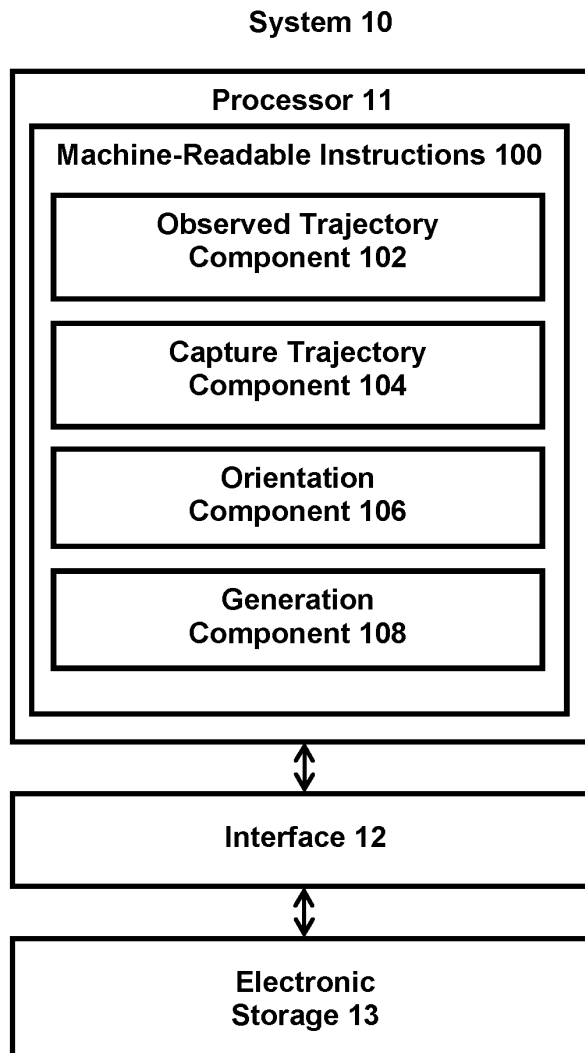


FIG. 1

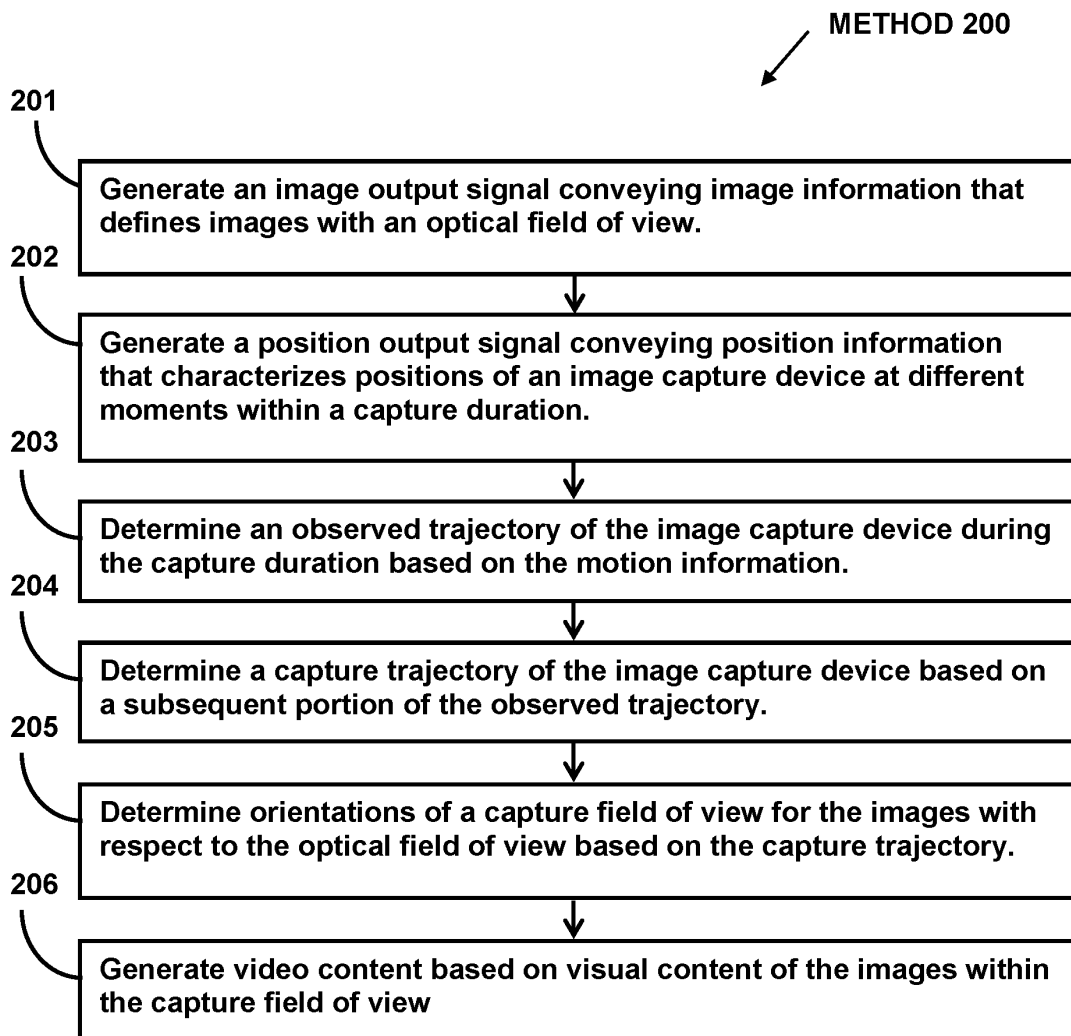


FIG. 2

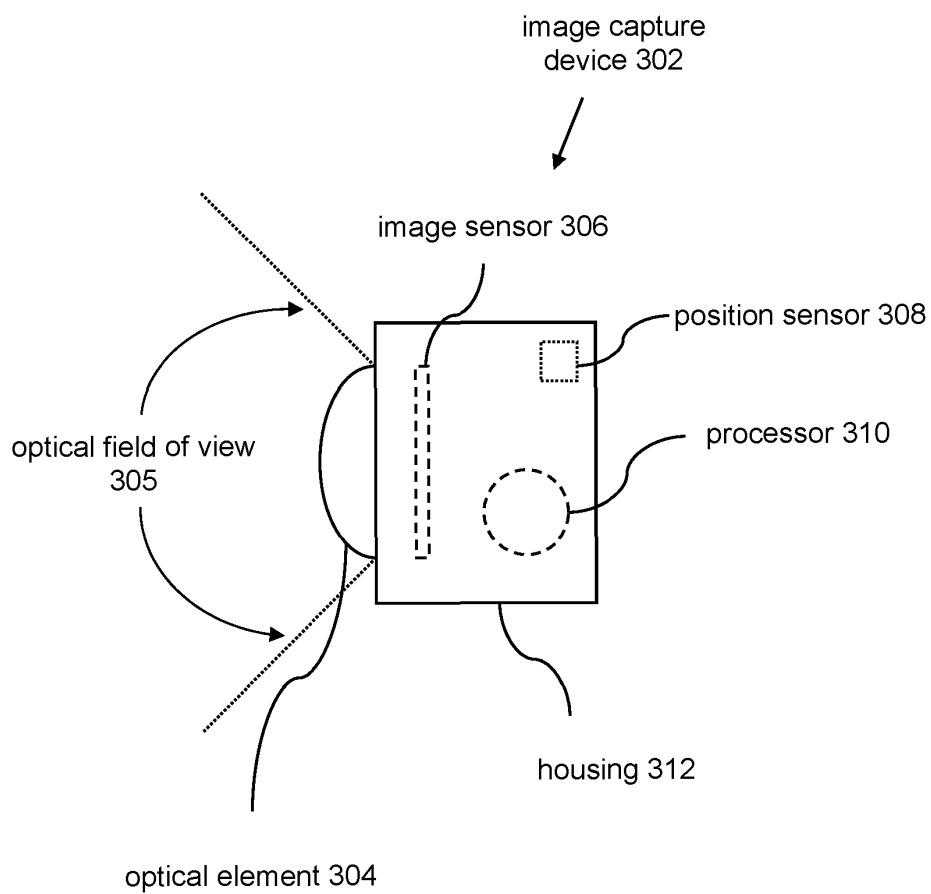


FIG. 3

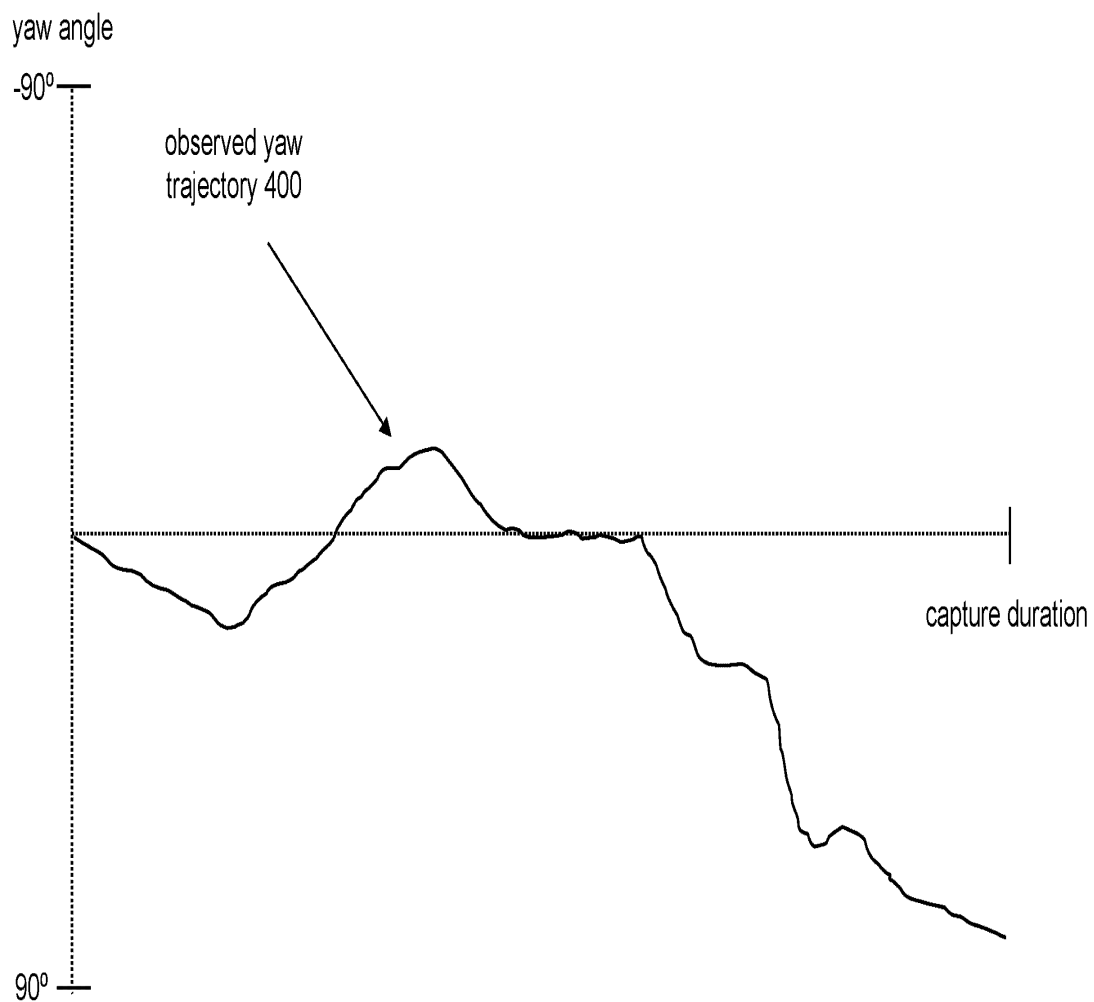


FIG. 4

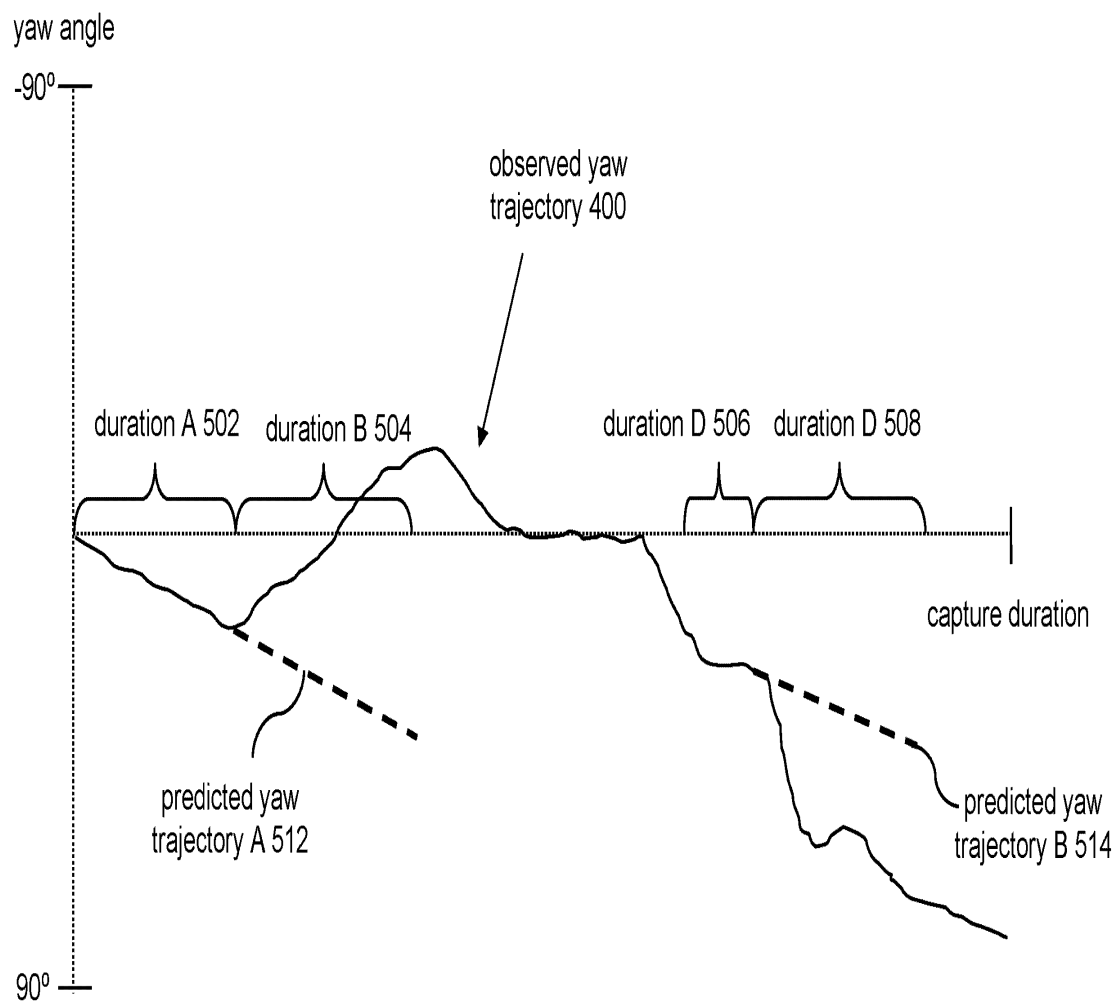


FIG. 5A

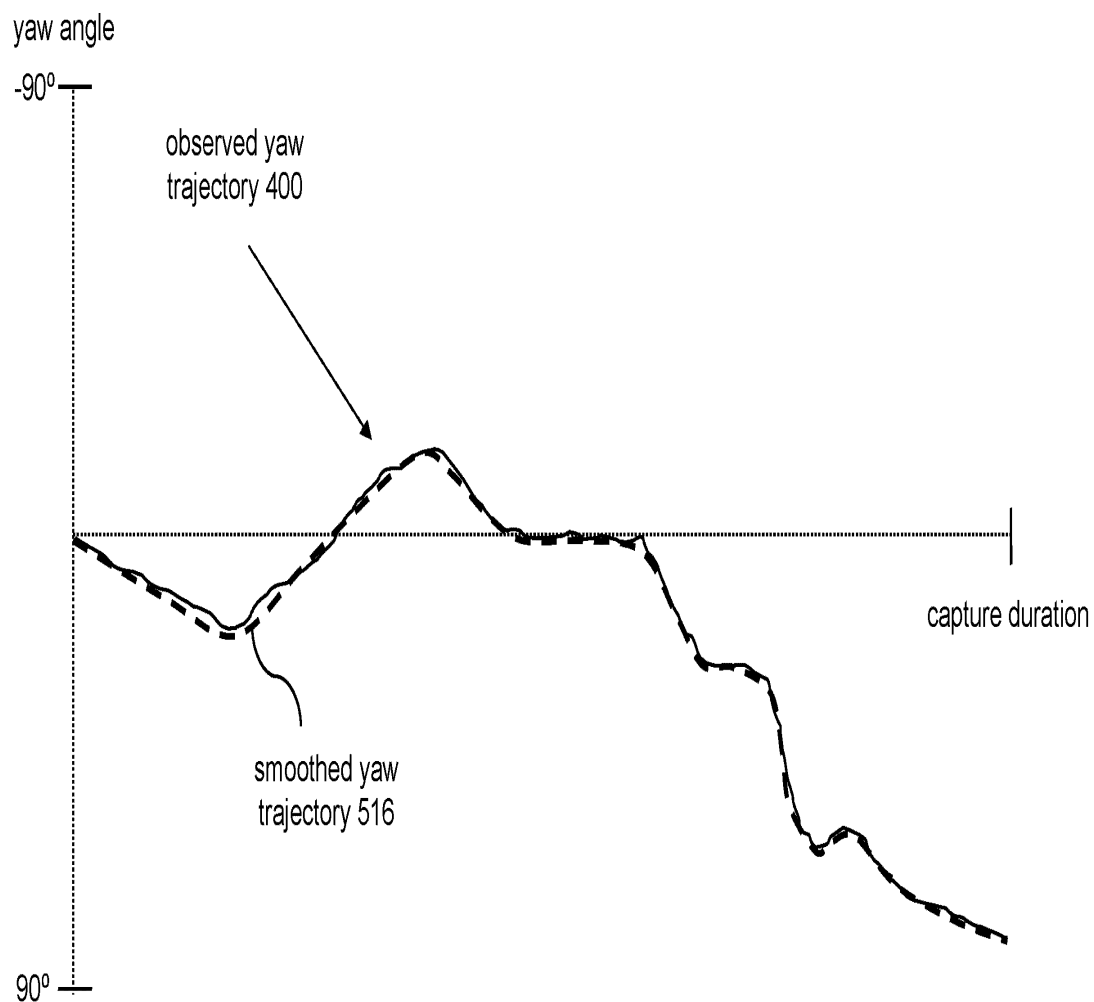


FIG. 5B

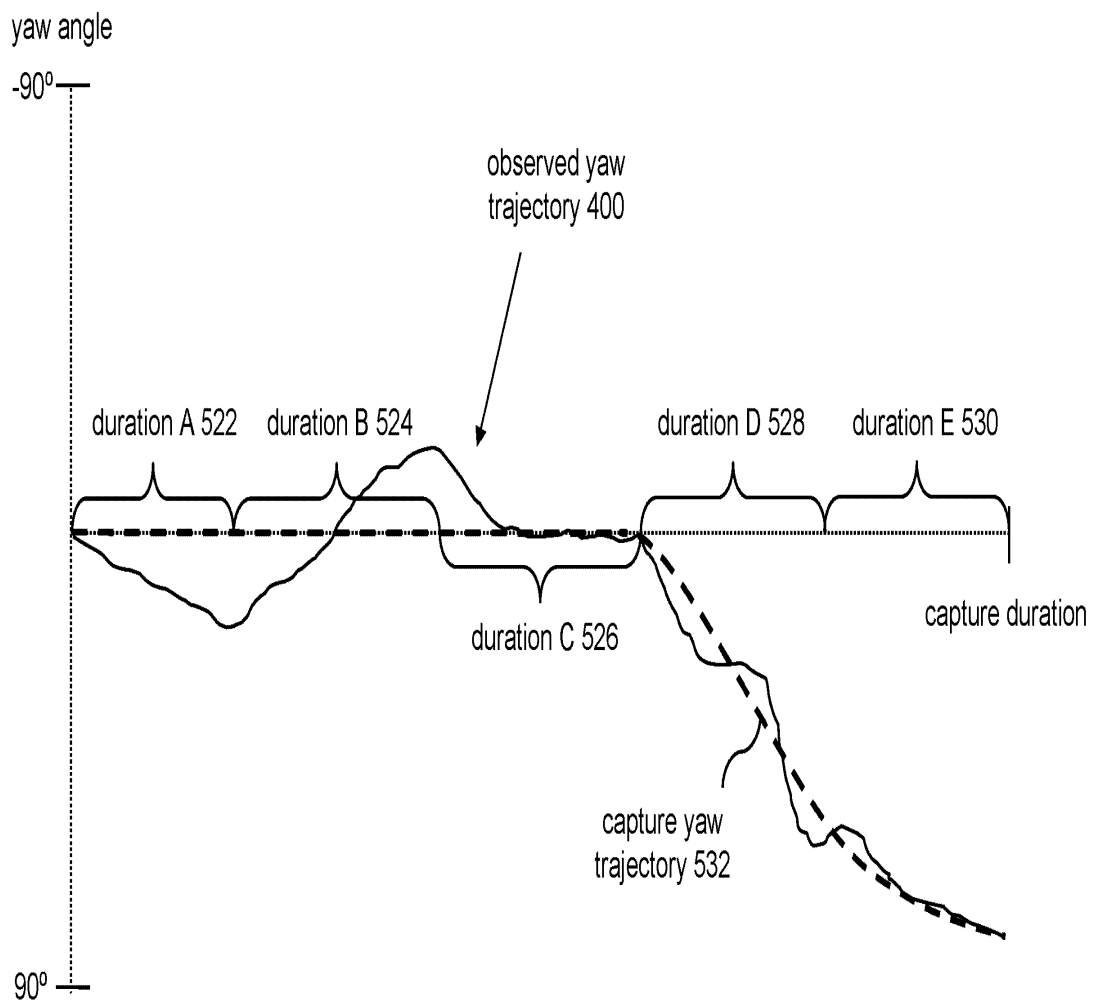


FIG. 5C

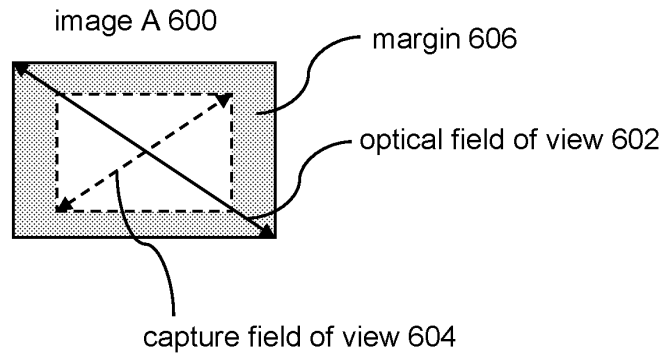


FIG. 6A

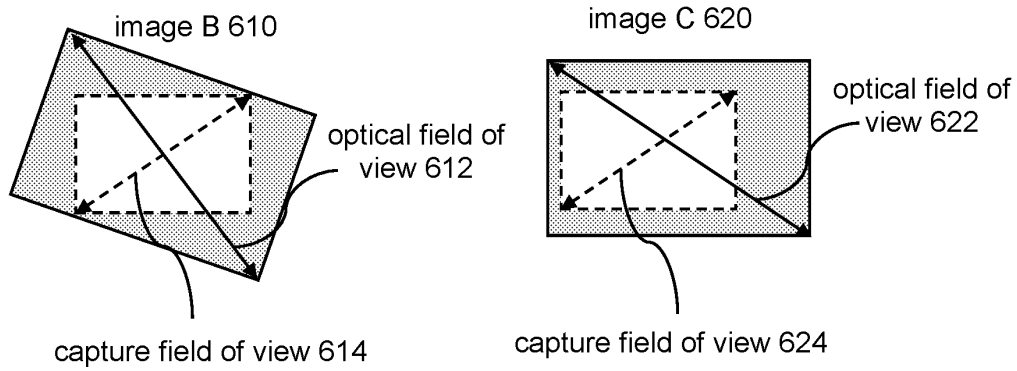


FIG. 6B

Appendix B

SYSTEMS AND METHODS FOR STABILIZING VIEWS OF VIDEOS***FIELD***

(01) This disclosure relates to stabilizing views of videos by stabilizing changes in viewing directions.

BACKGROUND

(02) Viewing directions for a video may include abrupt changes that result in a jerky/shaky view of the video. Such views of the video may decrease the enjoyability of viewing the video.

SUMMARY

(03) This disclosure relates to stabilizing views of videos. Video information, viewing information for the spherical video content, margin information, and/or other information may be obtained. The video information may define spherical video content. The spherical video content may have a progress length and include spherical video frames that define visual content viewable from a point of view as a function of progress through the progress length. The viewing information for the spherical video content may define a trajectory of viewing directions for the spherical video content. The trajectory of viewing direction may include viewing directions from the point of view as the function of progress through the progress length. The margin information may define a margin constraint that limits deviations from the trajectory of viewing directions. Stabilized viewing information may be determined based on the viewing information, the

SMRH:486594757.1

-1-

46DG-274118

margin information, and/or other information. The stabilized viewing information may define a stabilized trajectory of viewing directions for the spherical video content. The stabilized trajectory of viewing directions may include stabilized viewing directions from the point of view as the function of progress through the progress length. Differences between the trajectory of viewing directions and the stabilized trajectory of viewing directions may be limited by the margin constraint and/or other information. The spherical video content may be presented on a display based on the stabilized viewing information and/or other information.

(04) A system that stabilizes views of videos may include one or more electronic storages, one or more processors, and/or other components. An electronic storage may store video information, information relating to video content, viewing information, information relating to viewing directions, margin information, information relating to margin constraint, stabilized viewing information, information relating to stabilized viewing directions, and/or other information.

(05) The processor(s) may be configured by machine-readable instructions. Executing the machine-readable instructions may cause the processor(s) to facilitate stabilizing views of videos. The machine-readable instructions may include one or more computer program components. The computer program components may include one or more of a video information component, a viewing information component, a margin information component, a stabilization component, a presentation component, and/or other computer program components.

(06) The video information component may be configured to obtain video information and/or other information. The video information may define spherical video content. The spherical video content may have a progress length. The spherical video content may include spherical video frames that define visual content viewable from a point of view as a function of progress through the progress length.

(07) The viewing information component may be configured to obtain viewing information for the spherical video content and/or other information. The viewing information may define one or more trajectories of viewing directions for the spherical video content. A trajectory of viewing direction may include viewing directions from the point of view as the function of progress through the progress length.

(08) In some implementations, the viewing directions from the point of view may be defined based on rotations about the point of view. The rotations about the point of view may include rotations about one or more axes running through the point of view. The axe(s) may include a yaw axis, a pitch axis, a roll axis, and/or other axes.

(09) In some implementations, a trajectory of viewing directions may be defined based on a user's interaction with a display during presentation of the spherical video content on the display and/or other information.

(10) The margin information component may be configured to obtain margin information and/or other information. The margin information may define one or more margin constraints. A margin constraint may limit deviations from a trajectory of viewing directions. In some implementations, the margin constraint may include a yaw axis

constraint, a pitch axis constraint, a roll axis constraint, and/or other constraints. In some implementations, the margin constraint may include a target constraint and/or other constraints. The target constraint may be determined based on positions of a target in the spherical video frames and/or other information.

(11) The stabilization component may be configured to determine stabilized viewing information based on the viewing information, the margin information, and/or other information. The stabilized viewing information may define one or more stabilized trajectories of viewing directions for the spherical video content. A stabilized trajectory of viewing directions may include stabilized viewing directions from the point of view as the function of progress through the progress length. Differences between a trajectory of viewing directions and a stabilized trajectory of viewing directions may be limited by the margin constraint(s). Changes in the stabilized viewing directions may be smoother than changes in the viewing directions.

(12) In some implementations, determination of the stabilized viewing information based on the viewing information may include determining at least a portion of the stabilized trajectory of viewing directions based on a subsequent portion of the trajectory of viewing directions and/or other information. For example, the trajectory of viewing directions may include a first portion corresponding to a first moment within the progress length and a second portion corresponding to a second moment within the progress length. The second moment may be subsequent to the first moment. A portion of the stabilized trajectory of viewing directions corresponding to the first portion

of the trajectory of viewing directions may be determined based on the second portion of the trajectory of viewing directions and/or other information.

(13) The presentation component may be configured to present the spherical video content on a display based on the stabilized viewing information and/or other information.

(14) These and other objects, features, and characteristics of the system and/or method disclosed herein, as well as the methods of operation and functions of the related elements of structure and the combination of parts and economies of manufacture, will become more apparent upon consideration of the following description and the appended claims with reference to the accompanying drawings, all of which form a part of this specification, wherein like reference numerals designate corresponding parts in the various figures. It is to be expressly understood, however, that the drawings are for the purpose of illustration and description only and are not intended as a definition of the limits of the invention. As used in the specification and in the claims, the singular form of “a,” “an,” and “the” include plural referents unless the context clearly dictates otherwise.

BRIEF DESCRIPTION OF THE DRAWINGS

(15) FIG. 1 illustrates a system that stabilizes views of videos.

(16) FIG. 2 illustrates a method for stabilizing views of videos.

(17) FIG. 3 illustrates an example spherical video content.

(18) FIG. 4 illustrates example viewing directions for spherical video content.

- (19) FIGS. 5A-5B illustrate example extents of spherical video content.
- (20) FIG. 6 illustrates an example mobile device for determining viewing information for spherical video content.
- (21) FIG. 7 illustrates an example trajectory of viewing directions for video content.
- (22) FIG. 8 illustrates an example margin constraint.
- (23) FIG. 9 illustrates example trajectory of viewing direction and stabilized trajectory of viewing directions.

DETAILED DESCRIPTION

(24) FIG. 1 illustrates a system 10 for stabilizing views of videos. The system 10 may include one or more of a processor 11, an interface 12 (e.g., bus, wireless interface), an electronic storage 13, a display 14, and/or other components. Video information, viewing information for the spherical video content, margin information, and/or other information may be obtained by the processor 11. The video information may define spherical video content. The spherical video content may have a progress length and include spherical video frames that define visual content viewable from a point of view as a function of progress through the progress length. The viewing information for the spherical video content may define a trajectory of viewing directions for the spherical video content. The trajectory of viewing direction may include viewing directions from the point of view as the function of progress through the progress length. The margin information may define a margin constraint that limits deviations from the trajectory of viewing directions.

(25) Stabilized viewing information may be determined by the processor based on the viewing information, the margin information, and/or other information. The stabilized viewing information may define a stabilized trajectory of viewing directions for the spherical video content. The stabilized trajectory of viewing directions may include stabilized viewing directions from the point of view as the function of progress through the progress length. Differences between the trajectory of viewing directions and the stabilized trajectory of viewing directions may be limited by the margin constraint and/or other information. The spherical video content may be presented on the display 14 based on the stabilized viewing information and/or other information.

(26) The display 14 may include one or more devices that may present information visually. The display 14 may be configured to present video content (e.g., spherical video content) and/or other information. The display 14 may include one or more screens for presenting video content. For example, the display 14 may include a single screen in which the video content may be presented. As another example, the display 14 may include multiple screens in which the video content may be presented, with individual screens presenting portions of the video content. In some implementations, the display 14 may include a head-mounted display and the video content may be presented on the head-mounted display as virtual reality content. In some implementations, the display 14 may include a display of a mobile device (e.g., camera, smartphone, tablet, laptop).

(27) The electronic storage 13 may be configured to include electronic storage medium that electronically stores information. The electronic storage 13 may store

software algorithms, information determined by the processor 11, information received remotely, and/or other information that enables the system 10 to function properly. For example, the electronic storage 13 may store information relating to video information, video content (e.g., spherical video content), information relating to video content, viewing information, information relating to viewing directions, margin information, information relating to margin constraint, stabilized viewing information, information relating to stabilized viewing directions, and/or other information.

(28) Video content may refer to media content that may be consumed as one or more videos/video clips. Video content may include one or more videos/video clips stored in one or more formats/containers, and/or other video content. A format may refer to one or more ways in which the information defining video content may be arranged/laid out (e.g., file format). A container may refer to one or more ways in which information defining video content may be arranged/laid out in association with other information (e.g., wrapper format). Video content may include a video clip captured by an image capture device, multiple video clips captured by an image capture device, and/or multiple video clips captured by different image capture devices. Video content may include multiple video clips captured at the same time and/or multiple video clips captured at different times. Video content may include a video clip processed by an image/video application, multiple video clips processed by an image/video application, and/or multiple video clips processed by different image/video applications.

(29) Video content may have a progress length. A progress length may be defined in terms of time durations and/or frame numbers. For example, video content may include

a video having a time duration of 60 seconds. Video content may include a video having 1800 video frames. Video content having 1800 video frames may have a play time duration of 60 seconds when viewed at 30 frames/second. Other progress lengths, time durations, and frame numbers are contemplated.

(30) Video content may define visual content viewable as a function of progress through the progress length of the video content. Visual content of the video content may be included within video frames of the video content. That is, video content may include video frames that define visual content of the video content. In some implementations, video content may include one or more spherical video content, virtual reality content, and/or other video content. Spherical video content and/or virtual reality content may define visual content viewable from a point of view as a function of progress through the progress length of the spherical video/virtual reality content. Spherical video content may include spherical video frames that define visual content viewable from a point of view as a function of progress through the progress length of the spherical video content.

(31) Spherical video content may refer to a video capture of multiple views from a location. Spherical video content may include a full spherical video capture (360-degrees of capture, including opposite poles) or a partial spherical video capture (less than 360-degrees of capture). Spherical video content may be captured through the use of one or more image capture devices (e.g., cameras, image sensors) to capture images/videos from a location. Spherical video content may be generated based on light received within a field of view of a single image sensor or within fields of view of

multiple image sensors during a capture period. For example, multiple images/videos captured by multiple cameras/image sensors may be combined/stitched together to form the spherical video content. The field of view of camera(s)/image sensor(s) may be moved/rotated (*e.g.*, via movement/rotation of optical element(s), such as lens, of the image sensor(s)) to capture multiple images/videos from a location, which may be combined/stitched together to form the spherical video content.

(32) Visual content of the spherical video content may be included within spherical video frames of the spherical video content. A spherical video frame may include a spherical image of the spherical video content at a moment within the progress length of the spherical video content. For example, multiple images captured by multiple cameras/images sensors at a moment in time may be combined/stitched together to form a spherical video frame for the moment in time. A spherical video frame may include a full spherical image capture (360-degrees of capture, including opposite poles) or a particular spherical image capture (less than 360-degrees of capture). A spherical image (*e.g.*, spherical video frame) may be comprised of multiple sub-images (sub-frames). Sub-images may be generated by a single image sensor (*e.g.*, at different times as the field of view of the image sensor may be rotated) or by multiple image sensors (*e.g.*, individual sub-images for a moment in time captured by individual image sensors and combined/stitched together to form the spherical image).

(33) In some implementations, spherical video content may be stored with a 5.2K resolution. Using a 5.2K spherical video content may enable viewing windows (*e.g.*, directed to a portion of a spherical video frame) for the spherical video content with

resolution close to 1080p. In some implementations, spherical video content may include 12-bit video frames. In some implementations, spherical video content may be consumed as virtual reality content.

(34) Virtual reality content may refer to content (e.g., spherical video content) that may be consumed via virtual reality experience. Virtual reality content may associate different directions within the virtual reality content with different viewing directions, and a user may view a particular direction within the virtual reality content by looking in a particular direction. For example, a user may use a virtual reality headset to change the user's direction of view. The user's direction of view may correspond to a particular direction of view within the virtual reality content. For example, a forward-looking direction of view for a user may correspond to a forward direction of view within the virtual reality content.

(35) Spherical video content and/or virtual reality content may have been captured at one or more locations. For example, spherical video content and/or virtual reality content may have been captured from a stationary position (e.g., a seat in a stadium). Spherical video content and/or virtual reality content may have been captured from a moving position (e.g., a moving bike). Spherical video content and/or virtual reality content may include video capture from a path taken by the capturing device(s) in the moving position. For example, spherical video content and/or virtual reality content may include video capture from a person walking around in a music festival.

(36) FIG. 3 illustrates an example video content 300 defined by video information.

The video content 300 may include spherical video content. The video content 300 may

define visual content viewable from a point of view (e.g., center of sphere) as a function of progress through the progress length of the video content 300. FIG. 3 illustrates example rotational axes for the video content 300. Rotational axes for the video content 300 may include a yaw axis 310, a pitch axis 320, a roll axis 330, and/or other axes. Rotations about one or more of the yaw axis 310, the pitch axis 320, the roll axis 330, and/or other axes may define viewing directions/viewing window for the video content 300.

(37) For example, a 0-degree rotation of the video content 300 around the yaw axis 310 may correspond to a front viewing direction. A 90-degree rotation of the video content 300 around the yaw axis 310 may correspond to a right viewing direction. A 180-degree rotation of the video content 300 around the yaw axis 310 may correspond to a back-viewing direction. A -90-degree rotation of the video content 300 around the yaw axis 310 may correspond to a left viewing direction.

(38) A 0-degree rotation of the video content 300 around the pitch axis 320 may correspond to a viewing direction that may be level with respect to horizon. A 45-degree rotation of the video content 300 around the pitch axis 320 may correspond to a viewing direction that may be pitched up with respect to horizon by 45-degrees. A 90-degree rotation of the video content 300 around the pitch axis 320 may correspond to a viewing direction that may be pitched up with respect to horizon by 90-degrees (looking up). A -45-degree rotation of the video content 300 around the pitch axis 320 may correspond to a viewing direction that may be pitched down with respect to horizon by 45-degrees. A -90-degree rotation of the video content 300 around the pitch axis 320

may correspond to a viewing direction that may be pitched down with respect to horizon by 90-degrees (looking down).

(39) A 0-degree rotation of the video content 300 around the roll axis 330 may correspond to a viewing direction that may be upright. A 90-degree rotation of the video content 300 around the roll axis 330 may correspond to a viewing direction that may be rotated to the right by 90-degrees. A -90-degree rotation of the video content 300 around the roll axis 330 may correspond to a viewing direction that may be rotated to the left by 90-degrees. Other rotations and viewing directions are contemplated.

(40) A playback of video content (*e.g.*, the video content 300) may include presentation of one or more portions of the video content on one or more displays (*e.g.*, the display 14) based on a viewing window and/or other information. The viewing window may define extents of the visual content viewable on one or more displays as the function of progress through the progress length of the video content. The viewing window may define extents of the visual content presented on the display(s) as the function of progress through the progress length of the video content. For spherical video content, the viewing window may define extents of the visual content viewable from the point of view as the function of progress through the progress length of the spherical video content.

(41) The viewing window may be characterized by viewing directions, viewing sizes (*e.g.*, viewing zoom, viewing magnification), viewing rotations, and/or other information. A viewing direction may define a direction of view for video content. A viewing direction may define the angle/visual portion of the video content at which the viewing window

may be directed. A viewing direction may define a direction of view for the video content selected by a user, defined by instructions for viewing the video content, and/or determined based on other information about viewing the video content as a function of progress through the progress length of the video content (e.g., director track specifying viewing direction to be presented during playback as a function of progress through the progress length of the video content). For spherical video content, a viewing direction may define a direction of view from the point of view from which the visual content may be defined. Viewing directions for the video content may be characterized by rotations around the yaw axis 310, the pitch axis 320, the roll axis 330, and/or other axes. For example, a viewing direction of a 0-degree rotation of the video content around a yaw axis (e.g., the yaw axis 310) and a 0-degree rotation of the video content around a pitch axis (e.g., the pitch axis 320) may correspond to a front viewing direction (the viewing window may be directed to a forward portion of the visual content captured within the spherical video content).

(42) For example, FIG. 4 illustrates example changes in viewing directions 400 (e.g., selected by a user for video content, specified by a director track) as a function of progress through the progress length of the video content. The viewing directions 400 may change as a function of progress through the progress length of the video content. For example, at 0% progress mark, the viewing directions 400 may correspond to a zero-degree yaw angle and a zero-degree pitch angle. At 25% progress mark, the viewing directions 400 may correspond to a positive yaw angle and a negative pitch angle. At 50% progress mark, the viewing directions 400 may correspond to a zero-

degree yaw angle and a zero-degree pitch angle. At 75% progress mark, the viewing directions 400 may correspond to a negative yaw angle and a positive pitch angle. At 87.5% progress mark, the viewing directions 400 may correspond to a zero-degree yaw angle and a zero-degree pitch angle. Other viewing directions are contemplated.

(43) A viewing size may define a size (e.g., size, magnification, viewing angle) of viewable extents of visual content within the video content. A viewing size may define the dimensions of the viewing window. A viewing size may define a size of viewable extents of visual content within the video content selected by a user, defined by instructions for viewing the video content, and/or determined based on other information about viewing the video content as a function of progress through the progress length of the video content (e.g., director track specifying viewing size to be presented as a function of progress through the progress length of the video content). In some implementations, a viewing size may define different shapes of viewable extents. For example, a viewing window may be shaped as a rectangle, a triangle, a circle, and/or other shapes.

(44) A viewing rotation may define a rotation of the viewing window. A viewing rotation may define one or more rotations of the viewing window about one or more axis. In some implementations, a viewing rotation may be defined by one or more parameters of a viewing direction. For example, a viewing rotation may be defined based on rotation about an axis (e.g., the roll axis 330) corresponding to a viewing direction. A viewing rotation may define a rotation of the viewing window selected by a user, defined by instructions for viewing the video content, and/or determined based on

other information about viewing the video content as a function of progress through the progress length of the video content (e.g., director track specifying viewing rotation to be used as a function of progress through the progress length of the video content). For example, a viewing rotation of a viewing window having a rectangular shape may determine whether the rectangular viewing window is to be positioned in a portrait orientation (e.g., for a portrait view of the video content), in a landscape orientation (e.g., for a landscape view of the video content), and/or other orientation with respect to the visual content of the video content.

(45) FIGS. 5A-5B illustrate examples of extents for video content 500. In FIG. 5A, the size of the viewable extent of the video content 500 may correspond to the size of extent A 510. In FIG. 5B, the size of viewable extent of the video content 500 may correspond to the size of extent B 520. Viewable extent of the video content 500 in FIG. 5A may be smaller than viewable extent of the video content 500 in FIG. 5B. The viewable extent of the video content 500 in FIG. 5B may be more tilted with respect to the video content 500 than viewable extent of the video content 500 in FIG. 5A. Other viewing sizes and viewing rotations are contemplated.

(46) Referring to FIG. 1, the processor 11 may be configured to provide information processing capabilities in the system 10. As such, the processor 11 may comprise one or more of a digital processor, an analog processor, a digital circuit designed to process information, a central processing unit, a graphics processing unit, a microcontroller, an analog circuit designed to process information, a state machine, and/or other mechanisms for electronically processing information. The processor 11 may be

configured to execute one or more machine-readable instructions 100 to facilitate stabilizing views of videos. The machine-readable instructions 100 may include one or more computer program components. The machine-readable instructions 100 may include one or more of an a video information component 102, a viewing information component 104, a margin information component 106, a stabilization component 108, a presentation component 110, and/or other computer program components.

(47) The video information component 102 may be configured to obtain video information defining one or more video content (e.g., spherical video content) and/or other information. Obtaining video information may include one or more of accessing, acquiring, analyzing, determining, examining, identifying, loading, locating, opening, receiving, retrieving, reviewing, storing, and/or otherwise obtaining the video information. The video information component 102 may obtain video information from one or more locations. For example, the video information component 102 may obtain video information from a storage location, such as the electronic storage 13, electronic storage of information and/or signals generated by one or more image sensors, electronic storage of a device accessible via a network, and/or other locations. The video information component 102 may obtain video information from one or more hardware components (e.g., an image sensor) and/or one or more software components (e.g., software running on a computing device).

(48) The video information component 102 may be configured to obtain video information defining one or more video content during acquisition of the video content and/or after acquisition of the video content by one or more image sensors/image

capture devices. For example, the video information component 102 may obtain video information defining a video while the video is being captured by one or more image sensors/image capture devices. The video information component 102 may obtain video information defining a video after the video has been captured and stored in memory (e.g., the electronic storage 13).

(49) In some implementations, the video information component 102 may obtain video information based on user interaction with a user interface/application (e.g., video editing application), and/or other information. For example, a user interface/application may provide option(s) for a user to select one or more video content in which views are to be stabilized. The video information defining the video content may be obtained based on the user's selection of the video content through the user interface/video application. Other selections of video content are contemplated.

(50) In some implementations, one or more corrections may be applied to the video content. A correction may refer to a process that changes one or more visual and/or audio characteristics of the video content. A correction may be applied to the entire progress length of the video content, a portion of the progress length of the video content, the entire visual extent of the video content, or a portion of the visual extent of the video content. For example, a rotation/horizontal correction may be applied to video content to stabilize the video content. For instance, the video content may have been captured while the image capture device(s) capturing the video content was rotated with respect to the horizon. Such capture of the video content may result in a view of the video content showing rotated (e.g., tilted) view of the captured content. The video

content may be corrected for such rotation during capture such that the video content is stabilized about the horizon.

(51) As another example, a rolling shutter correction may be applied to video content. For instance, the video content may have been captured using rolling shutter where pixels of the video frames of the video content are captured on a row-by-row basis. Such capture of video content may result in one or more rolling shutter effects in the video content, such as wobble, skew, spatial aliasing, temporal aliasing, and/or other rolling shutter effects. The video content may be corrected for rolling shutter capture such that one or more rolling shutter effects are removed from the video content. Application of other corrections to the video content are contemplated.

(52) The viewing information component 104 may be configured to obtain viewing information for the video content (e.g., spherical video content) and/or other information. Obtaining viewing information may include one or more of accessing, acquiring, analyzing, determining, examining, identifying, loading, locating, opening, receiving, retrieving, reviewing, storing, and/or otherwise obtaining the viewing information. The viewing information component 104 may obtain viewing information from one or more locations. For example, the viewing information component 104 may obtain viewing information from a storage location, such as the electronic storage 13, electronic storage of information and/or signals generated by one or more sensors, electronic storage of a device accessible via a network, and/or other locations. The viewing information component 104 may obtain viewing information from one or more hardware

components and/or one or more software components (e.g., software running on a computing device).

(53) The viewing information may define one or more trajectories of viewing directions for the video content (e.g., spherical video content). A trajectory of viewing directions may refer to a line, a path, and/or a progression that defines values of viewing directions for the video content as a function of progress through the progress length of the video content. That is, the trajectory of viewing directions may be formed by and/or define values of viewing directions to be used when the video content is viewed, such as values of viewing directions defined in a director track that includes information as to how the video content is to be presented on a display (e.g., the display 14). For instance, a trajectory of viewing direction may include viewing directions from a point of view of spherical video content as the function of progress through the progress length of the spherical video content.

(54) The viewing directions from a point of view for video content may be defined based on rotations about the point of view. The rotations about the point of view may include rotations about one or more axes running through the point of view. The axe(s) may include a yaw axis, a pitch axis, a roll axis, and/or other axes. For example, the axe(s) may include one or more of the yaw axis 310, the pitch axis 320, and/or the roll axis 330 shown in FIG. 3. An example viewing directions 400 defined based on rotations about multiple axes is shown in FIG. 4. Other trajectories of viewing directions are contemplated. In some implementations, the viewing information may define other values for one or more parameters of a viewing window for the video content, such as

viewing size and/or viewing rotation. That is, viewing information may define one or more trajectories of viewing sizes and/or viewing rotations for the video content. Thus, viewing information may define one or more parameters by which a viewing window for the video content may be determined. The viewing information may define one or more of the direction at which the video content is viewed (viewing direction), the extent of the visual content of the video content that is within the viewing window (viewing size), and/or the rotation of the viewing window (viewing rotation).

(55) In some implementations, a trajectory of viewing directions, viewing sizes, and/or viewing rotations may be defined based on a user's interaction with a display during presentation of the video content (e.g., spherical video content) on the display and/or other information. That is, a user may define parameters (e.g., viewing direction, viewing size, viewing rotation) of a viewing window for video content by interacting with the display during presentation of the video content. A user's interaction with the display may include the user's engagement with the display (e.g., user's engagement with a touchscreen display) and/or a user's engagement with a device including and/or coupled to the display (e.g., user's engagement with a mobile device including the display).

(56) For example, FIG. 6 illustrates an example mobile device 605. A trajectory of viewing directions, viewing sizes, and/or viewing rotations for video content may be defined based on a user's interaction with the mobile device 605 during presentation of the video content on the display of the mobile device 605. For example, the mobile device 605 may include one or more motion/rotation sensors (e.g., accelerometer,

gyroscope, inertial measurement unit, magnetometer) and the user's interaction with the display may include the user moving the mobile device 605 rotationally and/or transversally. For example, the user may move the mobile device 605 laterally to change the viewing direction for the video content in a lateral direction. The user may rotate the mobile device 605 about the yaw axis 610 and/or the pitch axis 620 to change the viewing direction for the video content in the direction of rotation (e.g., rotating to the right, rotating to the left, pitch-up, pitch-down). The user may rotate the mobile device 605 (e.g., tilt to the right, tilt to the left) to change the viewing rotation for the video content. The user may move the mobile device forward or back to change the viewing size for the video content (e.g., zoom in or zoom out).

(57) As another example, the mobile device 605 may include a touchscreen display, and the user may define a trajectory of viewing directions, viewing sizes, and/or viewing rotations for video content by interacting with the touchscreen display during presentation of the video content on the touchscreen display. The user's interaction with the touchscreen display may include the user's engagement with the touchscreen display to make one or more gestures on the touchscreen display. For instance, the user may make one or more panning gestures, one or more rotating gestures, one or more stretching gestures, one or more pinching gestures, and/or other gestures on the touchscreen display. The gesture(s) on the touchscreen display may define values for viewing directions, viewing sizes, and/or viewing rotations. For example, a panning gesture may change the viewing direction for the video content in a lateral direction (e.g., to the left, to the right, around yaw axis). The rotating gesture may change the

viewing rotation for the video content in a clockwise or anti-clockwise direction. The stretching gesture may change the viewing size for the video content by zooming in on the portion of the video content in the middle of the stretching gesture. The pinching gesture may change the viewing size for the video content by zooming out of the video content (e.g., zooming out from the portion of the video content in the middle of the pinching gesture).

(58) A trajectory of viewing directions, viewing sizes, and/or viewing rotations defined based on a user's interaction with a display during presentation of the video content may be used to create a custom view for the video content. That is, the values of the viewing directions, viewing sizes, and/or viewing rotations defined based on the user's interaction may be stored and used to present a particular view of the video content on playback. For example, a user's interaction with a display may define a trajectory of viewing directions 400 shown in FIG. 4. The values of the rotations about the yaw axis and the pitch axis in the viewing directions may be stored in a director track. The director track may be used to determine in what direction a viewing window for the video content should be directed when the video content is played. The director track may define other values for determining a viewing window for the video content, such as viewing sizes and/or viewing rotations. Different user interactions may be used to generate different director tracks. Different director tracks may be used to provide different views of the video content. For instance, different director tracks defining different viewing directions may provide views with different styles (e.g., dynamic style view, calm style view).

(59) Using a user's interaction with a display to define viewing directions, viewing sizes, and/or viewing rotations for video content may result in unsmooth views of the video content. For example, a user's movement/rotation of the mobile device 605 may result in high frequency noise in the viewing directions, such as based on the user's hand(s) holding the mobile device 605 shaking. As another example, a user's movement/rotation of the mobile device 605 may result in abrupt changes in the viewing directions that result in a jerky/shaky view of the video content. That is, due to unsmooth motion of the user in moving/rotating the mobile device 605, the changes in viewing directions for the video content may include stagger, jitter, and/or other jerk in change. Other user interactions are contemplated.

(60) For example, FIG. 7 illustrates an example trajectory of viewing directions 700 for video content. The trajectory of viewing directions 700 may be defined based on a user's interaction with a display during presentation of the video content on the display. The trajectory of viewing directions 700 may define values of yaw angle positions (e.g., rotational positions defined with respect to a yaw axis, rotation to the left or right) at which a viewing window for the video content was directed during playback of the video content. For example, the trajectory of viewing directions 700 may be defined based on the display being rotated in a negative yaw direction, rotated in a positive yaw direction, rotated near a zero-yaw direction, and then rotated in the negative yaw direction. Other types of trajectory of viewing directions are contemplated.

(61) Presenting a view of the video content based on the trajectory of viewing directions 700 may be undesirable. For example, presenting a view of the video content

based on the trajectory of viewing directions 700 may result in a view that is shaky and/or that appears to include unintended camera motion. For instance, sharp/quick changes in the yaw angle positions in the trajectory of viewing directions 700 may result in abrupt changes in the direction of visuals within the video (e.g., quick left or right virtual camera motion). Multiple changes in the yaw angle positions in the trajectory of viewing directions 700 may result in a view that is repeatedly changing the direction of view (e.g., to the right, to the left, to the front, to the right) and/or not providing a stabilized view of the video content.

(62) The margin information component 106 may be configured to obtain margin information and/or other information. Obtaining margin information may include one or more of accessing, acquiring, analyzing, determining, examining, identifying, loading, locating, opening, receiving, retrieving, reviewing, storing, and/or otherwise obtaining the margin information. The margin information component 106 may obtain margin information from one or more locations. For example, the margin information component 106 may obtain margin information from a storage location, such as the electronic storage 13, electronic storage of information and/or signals generated by one or more sensors, electronic storage of a device accessible via a network, and/or other locations. The margin information component 106 may obtain margin information from one or more hardware components and/or one or more software components (e.g., software running on a computing device).

(63) The margin information may define one or more margin constraints. A margin constraint may refer to one or more constraints by which a trajectory of viewing

directions, viewing sizes, and/or viewing rotations may be changed. A margin constraint may provide one or more ranges of margin with which a trajectory of viewing directions, viewing sizes, and/or viewing rotations may be changed. For example, a margin constraint may limit deviations from a trajectory of viewing directions, such as the trajectory of viewing directions 700 shown in FIG. 7. The margin constraint may limit how much the trajectory of viewing directions may be stabilized to provide a more stable view of the video content. That is, the margin constraint may limit deviations of a stabilized trajectory of viewing directions from the trajectory of viewing directions. A stabilized trajectory of viewing directions may include smoother/more stable changes in viewing directions than the trajectory of viewing directions from which it may be generated. A stabilized trajectory of viewing sizes and/or viewing rotations may include smoother/more stable changes in viewing sizes and/or viewing rotation than the trajectory from which it may be generated.

(64) The margin constraint may provide one or more margins to limit how much the values of the stabilized trajectory may differ from the values of the original trajectory. Such margin(s) may preserve the intention of the user in changing the parameters (e.g., viewing directions, viewing sizes, viewing rotations) for viewing video content while removing jitters, stagger, and/or other jerky motions (e.g., removing high-frequency vibrations).

(65) A margin constraint may provide the same or different deviation limits on the values of the trajectory of viewing directions, viewing sizes, and/or viewing rotations. For example, FIG. 8 illustrates an example margin constraint 800 including different

deviation limits for different parameters of a viewing window for video content. The margin constraint 800 may include a yaw axis constraint 802, a pitch axis constraint 804, a roll axis constraint 806, and a viewing size constraint 808. The yaw axis constraint 802 and the pitch axis constraint 804 may limit the amount of deviations with respect to the change in viewing directions for the stabilized trajectory of viewing directions. The roll axis constraint 806 may limit the amount of deviations with respect to the change in viewing rotations for the stabilized trajectory of viewing rotations (or a portion of the stabilized trajectory of viewing directions). The viewing size constraint 808 may limit the amount of deviations with respect to the change in viewing sizes for the stabilized trajectory of viewing sizes.

(66) The margin constraint 800 may provide for different deviation limits with the yaw axis constraint 802, the pitch axis constraint 804, the roll axis constraint 806, and the viewing size constraint 808. For example, determination of the stabilized trajectory of viewing directions based on the margin constraint 800 may allow for greater deviation in the roll of a viewing direction than the yaw of the viewing direction. Determination of the stabilized trajectory based on the margin constraint 800 may allow for greater changes of the viewing rotation than the viewing size of the viewing window. Use of such a margin constraint may allow for different stabilization of the viewing directions, viewing sizes, and/or viewing rotations. Higher margins may enable greater stabilization of a view for video content while lower margins may more closely preserve the user interaction that defined the original view for the video content. In some implementations, different margin constraints may be used to provide different styles of

view. For example, a margin constraint with low margins may be used to provide a very stabilized view of video content while a margin constraint with high margins may be used to provide a slightly stabilized view of video content. As another example, a margin constraint with different limits on different parameters may be provided to create different styles of view of video content, such as one view that greatly stabilizes viewing rolls while allow for large fluctuations in viewing sizes. In some implementations, different margin constraint may be used for different styles of view (e.g., active view, scenery view). In some implementations, different margin constraint may be used for different content captured within the video content (e.g., sporting event, music festival, third-person view capture, first-person view capture).

(67) In some implementations, the margin constraint may include a target constraint and/or other constraints. A target constraint may provide limit the amount of deviations with respect to the change in viewing directions, viewing sizes, and/or viewing rotations based on the presence of target within the video content. For example, the target constraint may be determined based on positions of a target in spherical video frames and/or other information such that a stabilized viewing window includes the target or a portion of the target. For example, the target constraint may define how far stabilized viewing directions may be changed from the original viewing directions such that the stabilized viewing windows stays directed at the target. The target constraint may define how much stabilized viewing sizes may be changed from the original viewing sizes such that the stabilized viewing window includes all of or certain portion(s) of the target and/or certain amount of area surrounding the target. The target constraint may

define how much stabilized viewing rotations may be changed from the original viewing rotations such that the stabilized viewing window stays level with the target or the environment of the target. Other constraints are contemplated.

(68) The stabilization component 108 may be configured to determine stabilized viewing information based on the viewing information, the margin information, and/or other information. The stabilized viewing information may define one or more stabilized trajectories of viewing directions, viewing sizes, and/or viewing rotations for the video content (e.g., spherical video content). The stabilization component 108 may determine the stabilized viewing information based on the viewing information to stabilize (e.g., smooth) the trajectory characterized by the viewing information. That is, the stabilized trajectory defined by the stabilized viewing information may be a stabilized version of the trajectory defined by the viewing information. For example, a stabilized trajectory of viewing directions may include stabilized viewing directions from the point of view as the function of progress through the progress length of video content (e.g., spherical video content). Differences between a trajectory of viewing directions and a stabilized trajectory of viewing directions may be limited by the margin constraint(s). Changes in the stabilized viewing directions may be smoother than changes in the viewing directions. For example, jitters in the trajectory of viewing directions, viewing sizes, and/or viewing rotations may be removed in the stabilized trajectory. As another example, an abrupt jump in the trajectory of viewing directions, viewing sizes, and/or viewing rotations may be replaced in the stabilized trajectory with a linear/curved ramp

to provide for a smoother changes in the viewing directions, viewing sizes, and/or viewing rotations.

(69) In some implementations, the stabilization component 108 may determine different stabilized viewing information based on different margin information (different margin constraints). For example, the stabilization component 108 may determine different stabilized trajectory of viewing directions based on different margin constraints to create different views/styles of views of the video content.

(70) FIG. 9 illustrates example trajectory of viewing direction 900 and stabilized trajectory of viewing directions 932. The trajectory of viewing directions 700 may define values of yaw angle positions (e.g., rotational positions defined with respect to a yaw axis, rotation to the left or right) at which a viewing window for the video content was directed during playback of the video content. The stabilized trajectory of viewing directions 932 may be determined such that changes in the stabilized viewing directions of the stabilized trajectory of viewing directions 932 are smoother than changes in the viewing directions of the trajectory of viewing directions 900.

(71) In some implementations, determination of the stabilized viewing information based on the viewing information may include determining at least a portion of the stabilized trajectory of viewing directions, viewing sizes, and/or viewing rotations based on a subsequent portion of the trajectory of viewing directions, viewing sizes, and/or viewing rotations (look ahead), and/or other information. For example, stabilizing a portion of a trajectory of viewing directions based on a subsequent portion of the trajectory of viewing directions may including stabilizing a portion of the viewing

directions corresponding to a given moment within the progress length based on one or more portions of the viewing directions corresponding to one or more subsequent moments (moment(s) past the given moment) within the progress length.

(72) A subsequent moment within the progress length may be adjacent to the given moment or not adjacent to the given moment. Using the subsequent portion(s) of the trajectory (look ahead) to stabilize the trajectory may enable the stabilization component 108 to determine stabilized viewing information that preserves a user's intended changes in the viewing window. The user's intended changes in the viewing window may refer to changes in the extents defined by the viewing window that the user planned/meant to carry out, such as intentional changes in the viewing directions, viewing sizes, and/or viewing rotations.

(73) For example, the trajectory of viewing directions 900 may include portions corresponding to different moments within the progress length, such as a duration A 922, a duration B 924, a duration C 926, a duration D 928, and a duration E 930. One or more portions of the trajectory of viewing directions 900 may be stabilized (e.g., as shown in the stabilized trajectory of viewing direction 932) based on subsequent portion(s) of the trajectory of viewing directions 900. For instance, one or more portions of the trajectory of viewing directions 900 in the "future" may be used to determine whether a particular change in the viewing directions defined by the trajectory of viewing directions 900 was an intended change or an unintended change. For example, changes in the trajectory of viewing directions 900 may include a negative change in yaw angle below zero-yaw degree during the duration A 922, followed by a positive

change in yaw angle past zero-yaw degree during the duration B 924, followed by a negative change in yaw angle to zero-yaw degree during the duration C 926. Such fluctuation of the trajectory of viewing directions 900 around the zero-yaw degree may indicate that the initial negative change in the viewing directions was not intended, that the positive change in the viewing directions was meant to correct for the preceding unintended motion, that the user overshot the correction to go past the zero-yaw degree before arriving at the zero-yaw degree.

(74) The stabilized trajectory of viewing directions 932 during the duration A 922 may be determined based on the portion of the trajectory of viewing directions 900 corresponding to the duration B 924 and/or the duration C 926, and the stabilized trajectory of viewing directions 932 during the duration B 924 may be determined based on the portion of the trajectory of viewing directions 900 corresponding to the duration C 926 such that the unintended motion during duration A 922 and the duration B 924 are removed in the stabilized trajectory of viewing directions 932. Similarly, based on a subsequent portion of the trajectory of viewing directions 900 corresponding to the duration E 930, the portion of the stabilized trajectory of viewing direction 932 for the duration D 928 may be determined to include smooth changes from the zero-yaw degree to negative yaw degree. Other stabilization of trajectory of viewing directions, viewing sizes, and/or viewing rotations based on a subsequent portion of the trajectory of viewing directions, viewing sizes, and/or viewing rotations are contemplated.

(75) The presentation component 110 may be configured to present the video content (e.g., spherical video content) on one or more displays (e.g., the display 14) based on

the stabilized viewing information and/or other information. For example, based on stabilized trajectory of viewing directions (such as shown in FIG. 9), particular extents of the video content may be presented on a display (e.g., the display 14). Such presentation of the video content may provide for a stabilized view of the video content.

(76) The presentation of the video content based on the stabilized viewing information may be provided as a “default” or “automatic” view for the video content. That is, rather than presenting the video content based on the viewing information, the video content may be presented based on the stabilized viewing information. Such presentation of the video content may make the users think and/or believe that they created the stabilized view of the video content, rather than that they created the (unsmoothed) view of the video content which was stabilized by the stabilization component 110. In some implementations, multiple views of the video content may be presented based on different stabilized viewing information. For example, different stabilized viewing information may define different stabilization of trajectory of viewing directions, viewing sizes, and/or viewing rotations to create different styles of view. The different styles of view of the video content may be presented for selection by a user.

(77) In some implementations, the presentation component 110 may be configured to generate one or more video edits of the video content. A video edit may refer to a particular arrangement and/or manipulation of one or more portions (e.g., video clips) of the video content. A video edit, such as a video summary, may include portion(s) of the video content to provide an overview of the video content. A video edit may include extents of the video content defined by stabilized trajectory of viewing directions,

viewing sizes, and/or viewing rotations. That is, the presentation component 110 may generate a video edit of the video content based on the stabilized viewing information.

(78) Stabilized viewing information and/or other information defining stabilized trajectory may be stored in one or more storage media. For example, the stabilized viewing information may be stored in the electronic storage 13, remote storage locations (storage media located at/accessible through a server), and/or other locations. In some implementations, the stabilized viewing information may be stored through one or more intermediary devices. For example, the processor 11 may be located within a computing device without a connection to the storage device (*e.g.*, the computing device lacks WiFi/cellular connection to the storage device). The stabilized viewing information may be stored through another device that has the necessary connection (*e.g.*, the computing device using a WiFi/cellular connection of a paired mobile device, such as a smartphone, tablet, laptop, to store information in one or more storage media). Other storage locations for and storage of the stabilized viewing information are contemplated.

(79) While the description herein may be directed to video content, one or more other implementations of the system/method described herein may be configured for other types media content. Other types of media content may include one or more of audio content (*e.g.*, music, podcasts, audiobooks, and/or other audio content), multimedia presentations, images, slideshows, visual content (one or more images and/or videos), and/or other media content.

(80) Implementations of the disclosure may be made in hardware, firmware, software, or any suitable combination thereof. Aspects of the disclosure may be implemented as instructions stored on a machine-readable medium, which may be read and executed by one or more processors. A machine-readable medium may include any mechanism for storing or transmitting information in a form readable by a machine (e.g., a computing device). For example, a tangible computer-readable storage medium may include read-only memory, random access memory, magnetic disk storage media, optical storage media, flash memory devices, and others, and a machine-readable transmission media may include forms of propagated signals, such as carrier waves, infrared signals, digital signals, and others. Firmware, software, routines, or instructions may be described herein in terms of specific exemplary aspects and implementations of the disclosure, and performing certain actions.

(81) In some implementations, some or all of the functionalities attributed herein to the system 10 may be provided by external resources not included in the system 10. External resources may include hosts/sources of information, computing, and/or processing and/or other providers of information, computing, and/or processing outside of the system 10.

(82) Although the processor 11, the electronic storage 13, and the display 14 are shown to be connected to the interface 12 in FIG. 1, any communication medium may be used to facilitate interaction between any components of the system 10. One or more components of the system 10 may communicate with each other through hard-wired communication, wireless communication, or both. For example, one or more

components of the system 10 may communicate with each other through a network. For example, the processor 11 may wirelessly communicate with the electronic storage 13. By way of non-limiting example, wireless communication may include one or more of radio communication, Bluetooth communication, Wi-Fi communication, cellular communication, infrared communication, or other wireless communication. Other types of communications are contemplated by the present disclosure.

(83) Although the processor 11 is shown in FIG. 1 as a single entity, this is for illustrative purposes only. In some implementations, the processor 11 may comprise a plurality of processing units. These processing units may be physically located within the same device, or the processor 11 may represent processing functionality of a plurality of devices operating in coordination. The processor 11 may be configured to execute one or more components by software; hardware; firmware; some combination of software, hardware, and/or firmware; and/or other mechanisms for configuring processing capabilities on the processor 11.

(84) It should be appreciated that although computer components are illustrated in FIG. 1 as being co-located within a single processing unit, in implementations in which the processor 11 comprises multiple processing units, one or more of computer program components may be located remotely from the other computer program components.

(85) While computer program components are described herein as being implemented via the processor 11 through machine-readable instructions 100, this is merely for ease of reference and is not meant to be limiting. In some implementations,

one or more functions of computer program components described herein may be implemented via hardware (e.g., dedicated chip, field-programmable gate array) rather than software. One or more functions of computer program components described herein may be software-implemented, hardware-implemented, or software and hardware-implemented

(86) The description of the functionality provided by the different computer program components described herein is for illustrative purposes, and is not intended to be limiting, as any of computer program components may provide more or less functionality than is described. For example, one or more of computer program components may be eliminated, and some or all of its functionality may be provided by other computer program components. As another example, the processor 11 may be configured to execute one or more additional computer program components that may perform some or all of the functionality attributed to one or more of computer program components described herein.

(87) The electronic storage media of the electronic storage 13 may be provided integrally (*i.e.*, substantially non-removable) with one or more components of the system 10 and/or removable storage that is connectable to one or more components of the system 10 via, for example, a port (e.g., a USB port, a Firewire port, *etc.*) or a drive (e.g., a disk drive, *etc.*). The electronic storage 13 may include one or more of optically readable storage media (e.g., optical disks, *etc.*), magnetically readable storage media (e.g., magnetic tape, magnetic hard drive, floppy drive, *etc.*), electrical charge-based storage media (e.g., EPROM, EEPROM, RAM, *etc.*), solid-state storage media (e.g.,

flash drive, *etc.*), and/or other electronically readable storage media. The electronic storage 13 may be a separate component within the system 10, or the electronic storage 13 may be provided integrally with one or more other components of the system 10 (e.g., the processor 11). Although the electronic storage 13 is shown in FIG. 1 as a single entity, this is for illustrative purposes only. In some implementations, the electronic storage 13 may comprise a plurality of storage units. These storage units may be physically located within the same device, or the electronic storage 13 may represent storage functionality of a plurality of devices operating in coordination.

(88) FIG. 2 illustrates method 200 for stabilizing views of videos. The operations of method 200 presented below are intended to be illustrative. In some implementations, method 200 may be accomplished with one or more additional operations not described, and/or without one or more of the operations discussed. In some implementations, two or more of the operations may occur substantially simultaneously.

(89) In some implementations, method 200 may be implemented in one or more processing devices (e.g., a digital processor, an analog processor, a digital circuit designed to process information, a central processing unit, a graphics processing unit, a microcontroller, an analog circuit designed to process information, a state machine, and/or other mechanisms for electronically processing information). The one or more processing devices may include one or more devices executing some or all of the operation of method 200 in response to instructions stored electronically on one or more electronic storage mediums. The one or more processing devices may include one or

more devices configured through hardware, firmware, and/or software to be specifically designed for execution of one or more of the operation of method 200.

(90) Referring to FIG. 2 and method 200, at operation 201, video information may be obtained. The video information may define spherical video content. The spherical video content may have a progress length. The spherical video content may include spherical video frames. The spherical video frames may define visual content viewable from a point of view as a function of progress through the progress length. In some implementation, operation 201 may be performed by a processor component the same as or similar to the video information component 102 (Shown in FIG. 1 and described herein).

(91) At operation 202, viewing information for the spherical video content may be obtained. The viewing information may defining a trajectory of viewing directions for the spherical video content. The trajectory of viewing direction may include viewing directions from the point of view as the function of progress through the progress length. In some implementation, operation 202 may be performed by a processor component the same as or similar to the viewing information component 104 (Shown in FIG. 1 and described herein).

(92) At operation 203, margin information may be obtained. The margin information may define a margin constraint that limits deviations from the trajectory of viewing directions. In some implementation, operation 203 may be performed by a processor component the same as or similar to the margin information component 106 (Shown in FIG. 1 and described herein).

(93) At operation 204, stabilized viewing information may be determined based on the viewing information and the margin information. The stabilized viewing information may define a stabilized trajectory of viewing directions for the spherical video content. The stabilized trajectory of viewing directions may include stabilized viewing directions from the point of view as the function of progress through the progress length. Differences between the trajectory of viewing directions and the stabilized trajectory of viewing directions may be limited by the margin constraint. In some implementation, operation 204 may be performed by a processor component the same as or similar to the stabilization component 108 (Shown in FIG. 1 and described herein).

(94) At operation 205, the spherical video content may be presented on a display based on the stabilized viewing information. In some implementation, operation 205 may be performed by a processor component the same as or similar to the presentation component 110 (Shown in FIG. 1 and described herein).

(95) Although the system(s) and/or method(s) of this disclosure have been described in detail for the purpose of illustration based on what is currently considered to be the most practical and preferred implementations, it is to be understood that such detail is solely for that purpose and that the disclosure is not limited to the disclosed implementations, but, on the contrary, is intended to cover modifications and equivalent arrangements that are within the spirit and scope of the appended claims. For example, it is to be understood that the present disclosure contemplates that, to the extent possible, one or more features of any implementation can be combined with one or more features of any other implementation.

What is claimed is:

1. A system that stabilizes views of videos, the system comprising:
one or more physical processors configured by machine-readable instructions to:
 obtain video information defining spherical video content, the spherical
 video content having a progress length and including spherical video frames that
 define visual content viewable from a point of view as a function of progress
 through the progress length;
 obtain viewing information for the spherical video content, the viewing
 information defining a trajectory of viewing directions for the spherical video
 content, the trajectory of viewing direction including viewing directions from the
 point of view as the function of progress through the progress length;
 obtain margin information, the margin information defining a margin
 constraint that limits deviations from the trajectory of viewing directions;
 determine stabilized viewing information based on the viewing information
 and the margin information, the stabilized viewing information defining a
 stabilized trajectory of viewing directions for the spherical video content, the
 stabilized trajectory of viewing directions including stabilized viewing directions
 from the point of view as the function of progress through the progress length,
 wherein differences between the trajectory of viewing directions and the
 stabilized trajectory of viewing directions are limited by the margin constraint; and
 present the spherical video content on a display based on the stabilized
 viewing information.

2. The system of claim 1, wherein the one or more physical processors are, to determine the stabilized viewing information based on the viewing information, further configured by the machine-readable instructions to determine at least a portion of the stabilized trajectory of viewing directions based on a subsequent portion of the trajectory of viewing directions.

3. The system of claim 2, wherein:

the trajectory of viewing directions includes a first portion corresponding to a first moment within the progress length and a second portion corresponding to a second moment within the progress length, the second moment subsequent to the first moment;
a portion of the stabilized trajectory of viewing directions corresponding to the first portion of the trajectory of viewing directions is determined based on the second portion of the trajectory of viewing directions.

4. The system of claim 1, wherein changes in the stabilized viewing directions are smoother than changes in the viewing directions.

5. The system of claim 1, wherein the viewing directions from the point of view are defined based on rotations about the point of view.

6. The system of claim 5, wherein the rotations about the point of view include rotations about one or more axes running through the point of view, the one or more axes including a yaw axis, a pitch axis, and/or a roll axis.
7. The system of claim 6, wherein the margin constraint includes a yaw axis constraint, a pitch axis constraint, and/or a roll axis constraint.
8. The system of claim 6, wherein the trajectory of viewing directions is defined based on a user's interaction with a display during presentation of the spherical video content on the display.
9. The system of claim 1, wherein the margin constraint includes a target constraint determined based on positions of a target in the spherical video frames.
10. A method for stabilizing views of videos, the method performed by a computing system including one or more processors, the method comprising:
- obtaining, by the computing system, video information defining spherical video content, the spherical video content having a progress length and including spherical video frames that define visual content viewable from a point of view as a function of progress through the progress length;
 - obtaining, by the computing system, viewing information for the spherical video content, the viewing information defining a trajectory of viewing directions for the

spherical video content, the trajectory of viewing direction including viewing directions from the point of view as the function of progress through the progress length;

obtaining, by the computing system, margin information, the margin information defining a margin constraint that limits deviations from the trajectory of viewing directions;

determining, by the computing system, stabilized viewing information based on the viewing information and the margin information, the stabilized viewing information defining a stabilized trajectory of viewing directions for the spherical video content, the stabilized trajectory of viewing directions including stabilized viewing directions from the point of view as the function of progress through the progress length, wherein differences between the trajectory of viewing directions and the stabilized trajectory of viewing directions are limited by the margin constraint; and

presenting, by the computing system, the spherical video content on a display based on the stabilized viewing information.

11. The method of claim 10, wherein determining the stabilized viewing information based on the viewing information includes determining at least a portion of the stabilized trajectory of viewing directions based on a subsequent portion of the trajectory of viewing directions.

12. The method of claim 11, wherein:

the trajectory of viewing directions includes a first portion corresponding to a first moment within the progress length and a second portion corresponding to a second moment within the progress length, the second moment subsequent to the first moment;

a portion of the stabilized trajectory of viewing directions corresponding to the first portion of the trajectory of viewing directions is determined based on the second portion of the trajectory of viewing directions.

13. The method of claim 10, wherein changes in the stabilized viewing directions are smoother than changes in the viewing directions.

14. The method of claim 10, wherein the viewing directions from the point of view are defined based on rotations about the point of view.

15. The method of claim 14, wherein the rotations about the point of view include rotations about one or more axes running through the point of view, the one or more axes including a yaw axis, a pitch axis, and/or a roll axis.

16. The method of claim 15, wherein the margin constraint includes a yaw axis constraint, a pitch axis constraint, and/or a roll axis constraint.

17. The method of claim 15, wherein the trajectory of viewing directions is defined based on a user's interaction with a display during presentation of the spherical video content on the display.

18. The method of claim 10, wherein the margin constraint includes a target constraint determined based on positions of a target in the spherical video frames.

19. A system that stabilizes views of videos, the system comprising:
one or more physical processors configured by machine-readable instructions to:
 obtain video information defining spherical video content, the spherical video content having a progress length and including spherical video frames that define visual content viewable from a point of view as a function of progress through the progress length;
 obtain viewing information for the spherical video content, the viewing information defining a trajectory of viewing directions for the spherical video content, the trajectory of viewing direction including viewing directions from the point of view as the function of progress through the progress length, wherein the viewing directions from the point of view are defined based on rotations about one or more axes running through the point of view, the one or more axes including a yaw axis, a pitch axis, and/or a roll axis;
 obtain margin information, the margin information defining a margin constraint that limits deviations from the trajectory of viewing directions, the

margin constraint including a yaw axis constraint, a pitch axis constraint, and/or a roll axis constraint;

determine a stabilized viewing information based on the viewing information and the margin information, the stabilized viewing information defining a stabilized trajectory of viewing directions for the spherical video content, the stabilized trajectory of viewing directions including stabilized viewing directions from the point of view as the function of progress through the progress length, wherein differences between the trajectory of viewing directions and the stabilized trajectory of viewing directions are limited by the margin constraint and changes in the stabilized viewing directions are smoother than changes in the viewing directions; and

present the spherical video content on a display based on the stabilized viewing information.

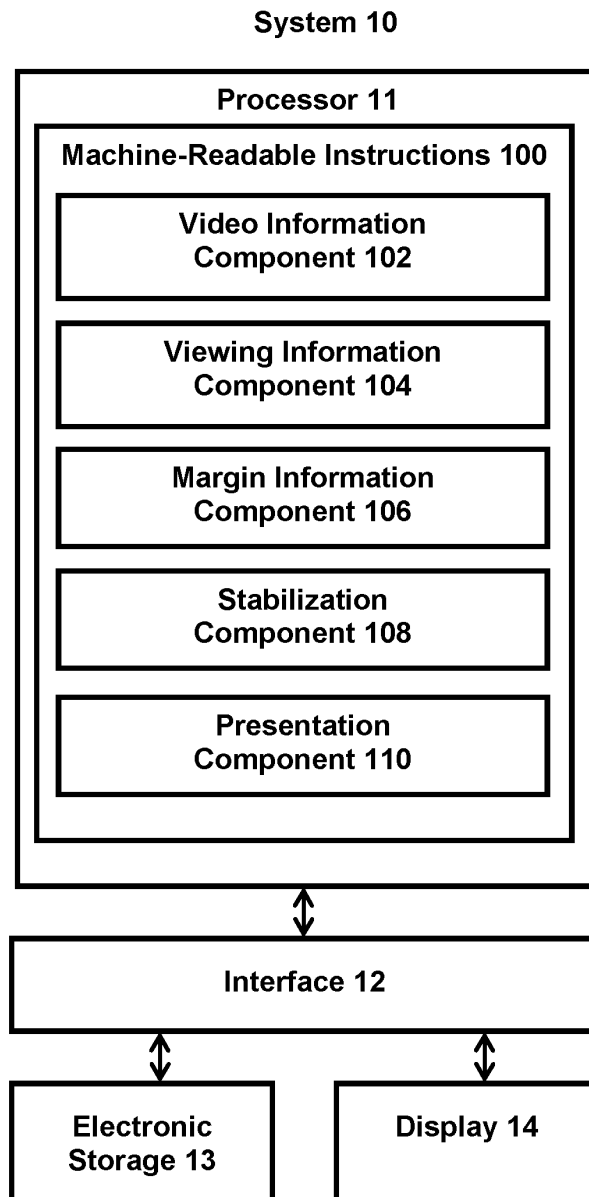
20. The system of claim 19, wherein the one or more physical processors are, to determine a stabilized viewing information based on the viewing information, further configured by the machine-readable instructions to determine at least a portion of the stabilized trajectory of viewing directions based on a subsequent portion of the trajectory of viewing directions, wherein:

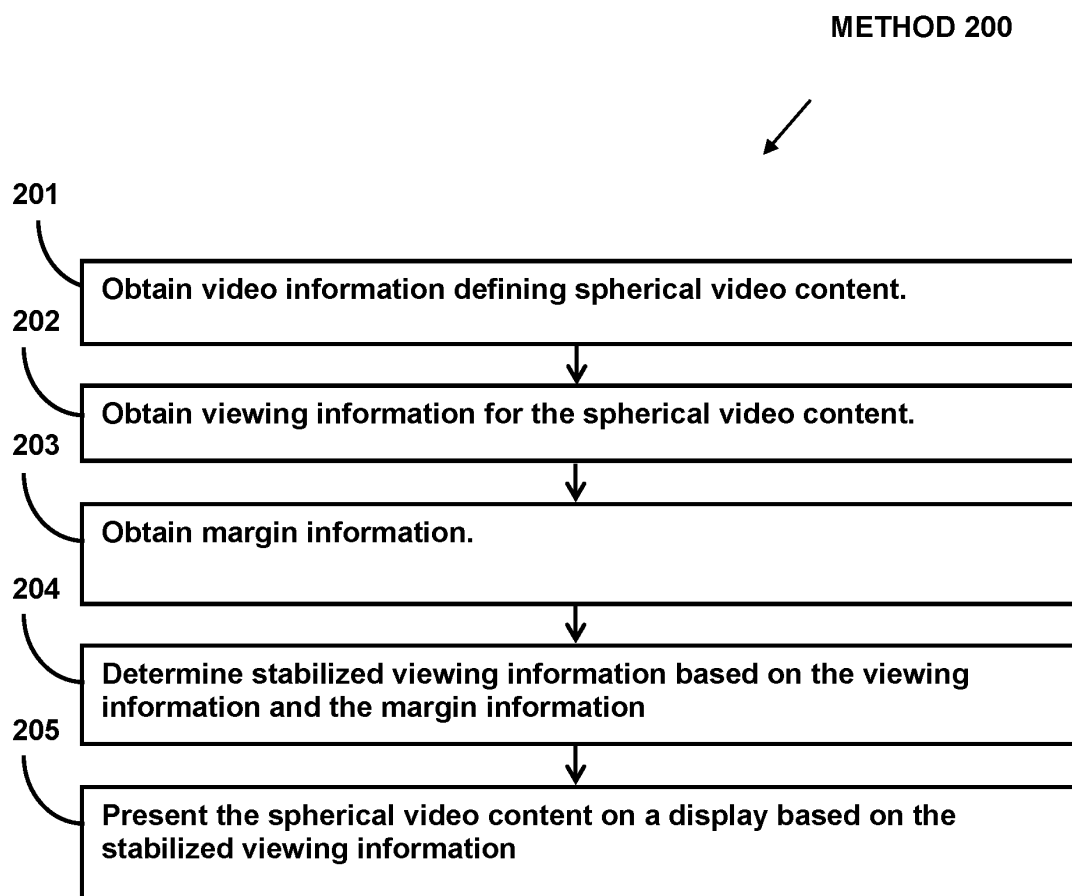
the trajectory of viewing directions includes a first portion corresponding to a first moment within the progress length and a second portion corresponding to a second moment within the progress length, the second moment subsequent to the first moment;

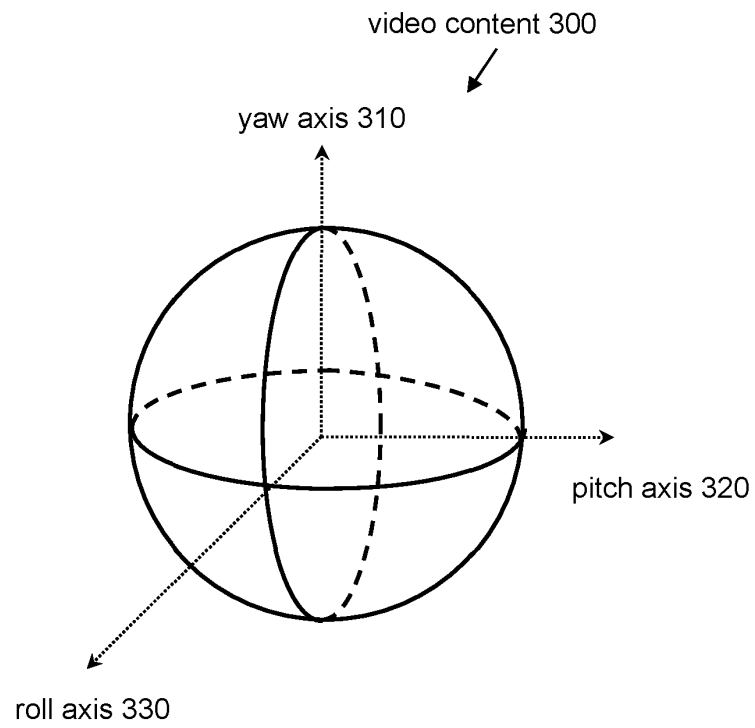
a portion of the stabilized trajectory of viewing directions corresponding to the first portion of the trajectory of viewing directions is determined based on the second portion of the trajectory of viewing directions.

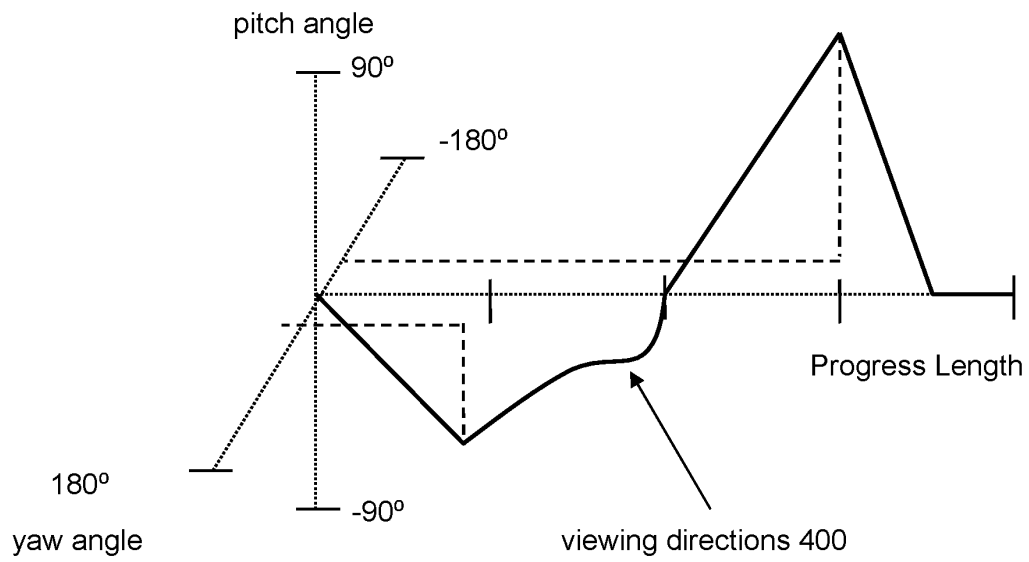
ABSTRACT

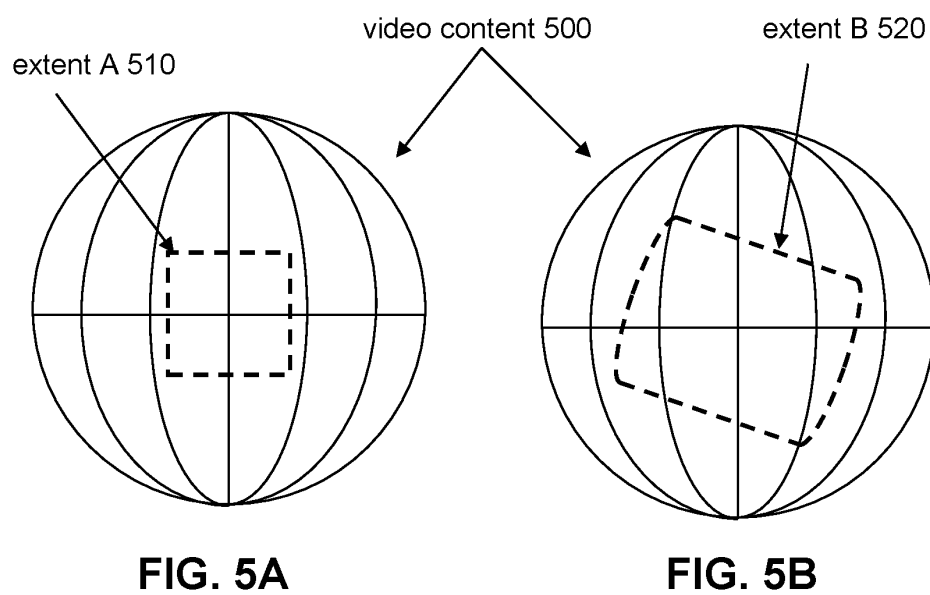
A viewing direction may define an angle/visual portion of a spherical video at which a viewing window is directed. A trajectory of viewing direction may include changes in viewing directions for playback of spherical video. Abrupt changes in the viewing directions may result in jerky or shaky views of the spherical video. Changes in the viewing directions may be stabilized to provide stabilized views of the spherical video. Amount of stabilization may be limited by a margin constraint.

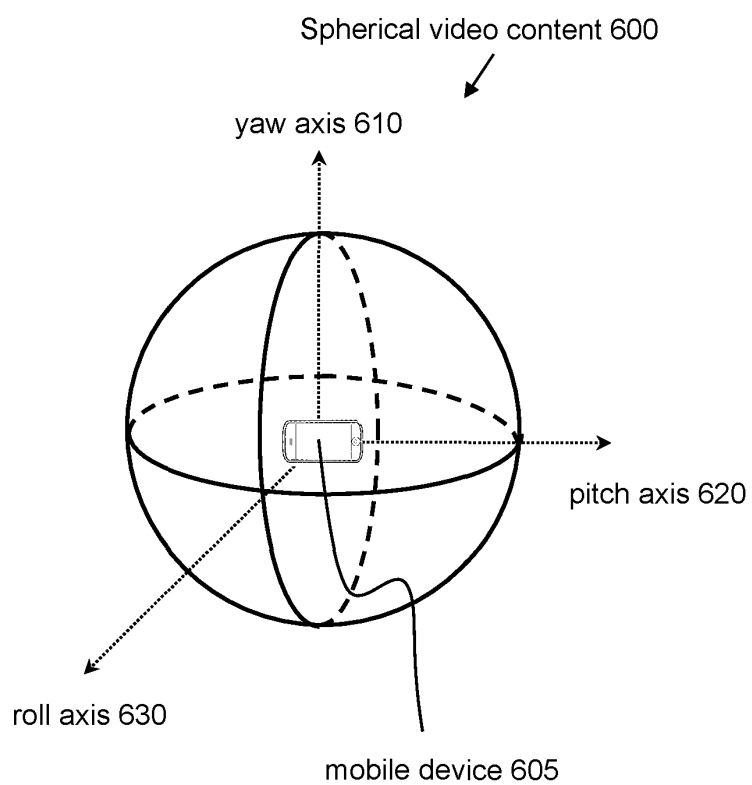
**FIG. 1**

**FIG. 2**

**FIG. 3**

**FIG. 4**



**FIG. 6**

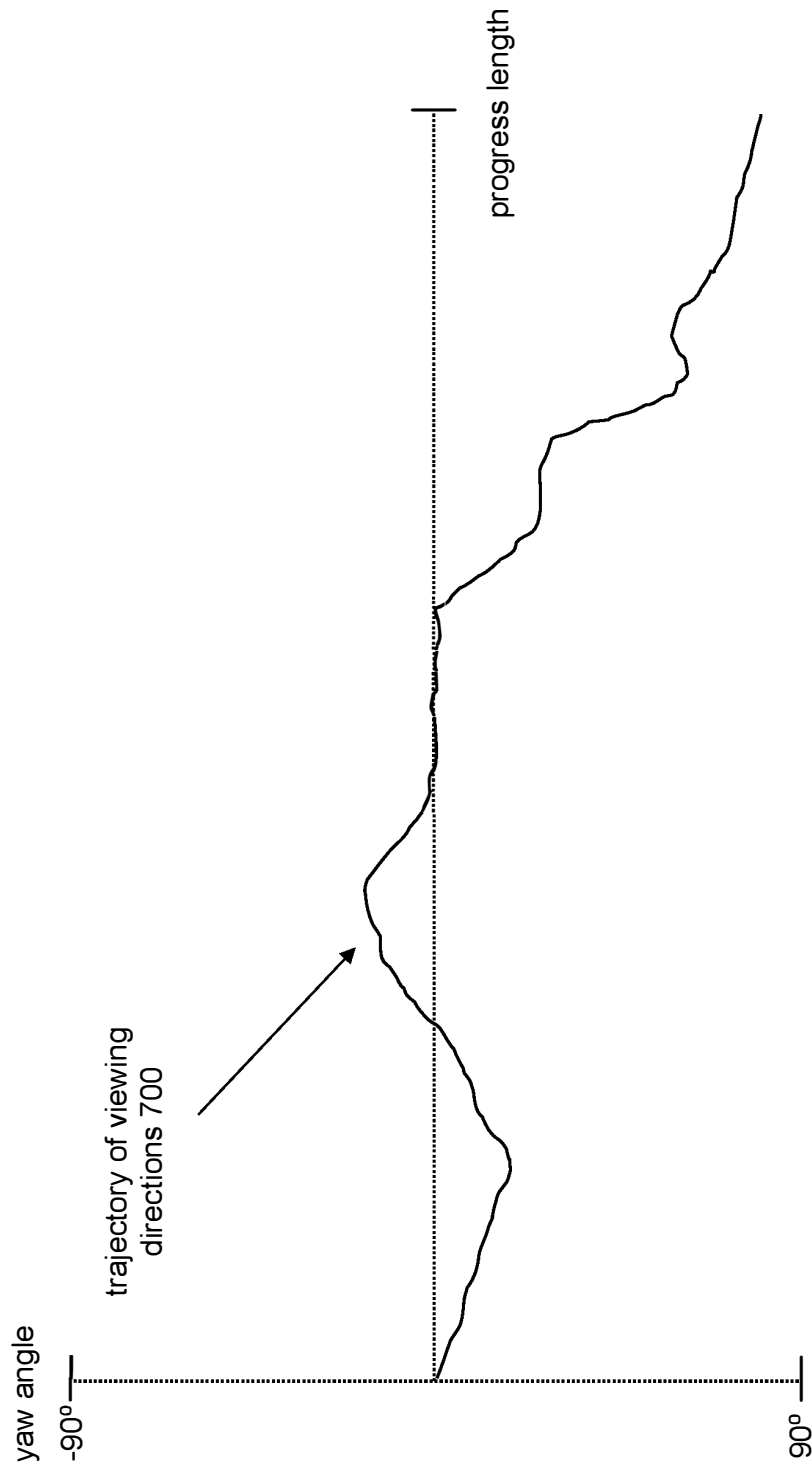
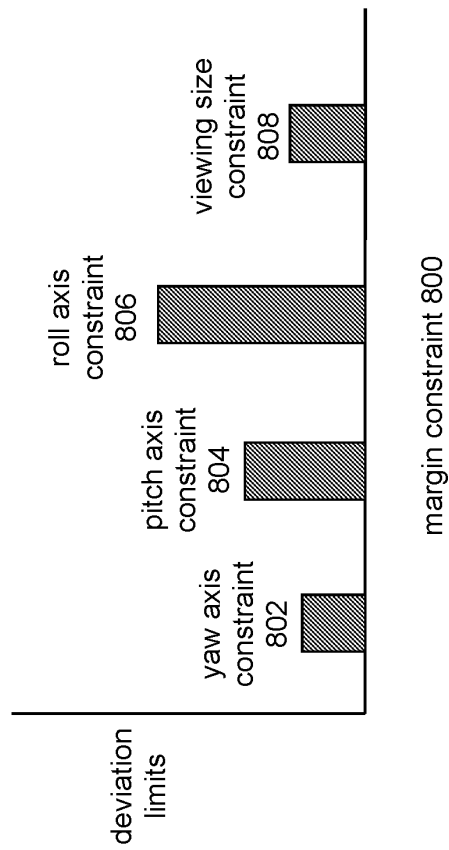


FIG. 7

**FIG. 8**

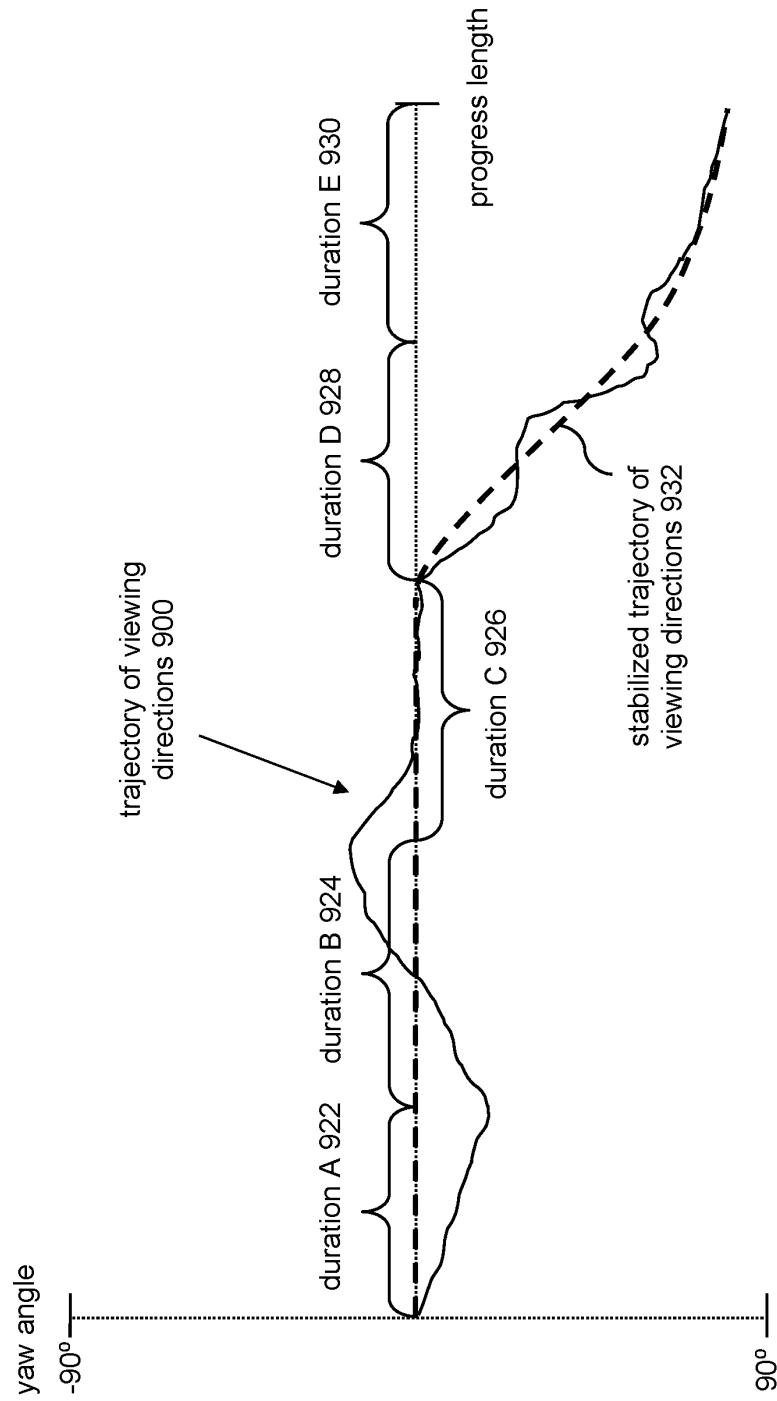


FIG. 9

Electronic Patent Application Fee Transmittal				
Application Number:				
Filing Date:				
Title of Invention:		SYSTEMS AND METHODS FOR STABILIZING VIDEOS		
First Named Inventor/Applicant Name:		César Douady		
Filer:		D. Benjamin Esplin/Chelsea Pearsall		
Attorney Docket Number:		46DG-286652		
Filed as Large Entity				
Filing Fees for Provisional				
Description	Fee Code	Quantity	Amount	Sub-Total in USD(\$)
Basic Filing:				
PROVISIONAL APPLICATION FILING	1005	1	280	280
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD(\$)
Extension-of-Time:				
Miscellaneous:				
Total in USD (\$)				280

Electronic Acknowledgement Receipt	
EFS ID:	33766230
Application Number:	62733237
International Application Number:	
Confirmation Number:	2932
Title of Invention:	SYSTEMS AND METHODS FOR STABILIZING VIDEOS
First Named Inventor/Applicant Name:	César Douady
Customer Number:	137002
Filer:	D. Benjamin Esplin/Chelsea Pearsall
Filer Authorized By:	D. Benjamin Esplin
Attorney Docket Number:	46DG-286652
Receipt Date:	19-SEP-2018
Filing Date:	
Time Stamp:	13:59:43
Application Type:	Provisional

Payment information:

Submitted with Payment	yes
Payment Type	DA
Payment was successfully received in RAM	\$280
RAM confirmation Number	092018INTEFSW00000251504561
Deposit Account	504561
Authorized User	Chelsea Pearsall
The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows: 37 CFR 1.16 (National application filing, search, and examination fees) 37 CFR 1.17 (Patent application and reexamination processing fees)	

37 CFR 1.19 (Document supply fees)					
37 CFR 1.20 (Post Issuance fees)					
37 CFR 1.21 (Miscellaneous fees and charges)					
File Listing:					
Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Application Data Sheet	WebADS.pdf	135332	no	8
			ed85c56cfa8ffef78b7bec5b87b8620625b77c40		
Warnings:					
Information:					
2		Specification.pdf	1695839	yes	127
			f7b6e301b16f424bdbc0f2b4ce8e722186fec25b		
	Multipart Description/PDF files in .zip description				
	Document Description		Start	End	
	Specification		1	13	
	Appendix to the Specification		14	68	
	Appendix to the Specification		69	127	
Warnings:					
Information:					
3	Fee Worksheet (SB06)	fee-info.pdf	30107	no	2
			f7e12460e521f909c114ef5f33cdf58c7a999dfc		
Warnings:					
Information:					
Total Files Size (in bytes):			1861278		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.